GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 01:54:29 ; Search time 615.146 Seconds
(without alignments)
1396.582 Million cell updates/sec

Title: US-09-692-077D-13
Perfect score: 21
Sequence: 1 gctcatcatccctttctcgct 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		21	100.0	21	6	AX350501	AX350501 Sequence
2		21	100.0	1149	4	AF337539	AF337539 Pteropus
3		21	100.0	1149	4	RAM315937	AJ315937 Rousettus
4		21	100.0	1150	4	CSP251181	AJ251181 Cynopteru
5		21	100.0	1156	4	MLA251107	AJ251107 Micropota
6		21	100.0	1162	4	FCA251174	AJ251174 Felis cat
7		21	100.0	1167	4	AF337538	AF337538 Hipposide
8		21	100.0	1167	4	NAL419805	AJ419805 Nyctimene
9		21	100.0	1167	4	RCR419806	AJ419806 Rhinoloph
10		21	100.0	1168	4	PVI251176	AJ251176 Phoca vit
11		21	100.0	1170	4	AF337543	AF337543 Taphozous
12		21	100.0	1171	4	DDAAR2B	Y15947 Dugong dugo
13		21	100.0	1171	4	TMA251109	AJ251109 Trichechu
14		21	100.0	1171	9	CVA251182	AJ251182 Cynocepha
15		21	100.0	1172	10	ARU427256	AJ427256 Aplodonti
16		21	100.0	1173	4	APA419813	AJ419813 Antrozous
17		21	100.0	1173	4	EAT419810	AJ419810 Emballonu
18		21	100.0	1174	4	PCA427417	AJ427417 Physeter
19		21	100.0	1176	4	AF337537	AF337537 Megaderma
20		21	100.0	1176	4	AF337542	AF337542 Tadarida
21		21	100.0	1176	4	LAL315940	AJ315940 Lagenorhy
22		21	100.0	1176	10	MMO427255	AJ427255 Marmota m
23		21	100.0	1176	10	SVU315942	AJ315942 Sciurus v
24		21	100.0	1177	4	BPH251175	AJ251175 Balaenopt
25		21	100.0	1177	4	BTAAR2B	Y15944 Bos taurus
26		21	100.0	1179	4	AF337540	AF337540 Myotis da
27		21	100.0	1180	4	LPA505821	AJ505821 Lama paco
28		21	100.0	1180	4	MCA251180	AJ251180 Macrotus
29		21	100.0	1180	9	NCO251186	AJ251186 Nycticebu
30		21	100.0	1185	10	ASP427259	AJ427259 Anomaluru
31		21	100.0	1185	10	GGL427258	AJ427258 Glis glis
32		21	100.0	1194	4	SCI315936	AJ315936 Sorex cin
33		21	100.0	1197	4	LPA315941	AJ315941 Lama paco
34		21	100.0	1197	4	MTE505820	AJ505820 Manis tet
35		21	100.0	1203	10	DME427261	AJ427261 Dipodomys
36		21	100.0	1344	6	AX350490	AX350490 Sequence
37		21	100.0	1344	9	AF316895	AF316895 Homo sapi
38		21	100.0	1353	6	AX350489	AX350489 Sequence
39		21	100.0	2072	6	AR270618	AR270618 Sequence
40		21	100.0	2072	9	HUMADRA2RA	M34041 Human alpha
41		21	100.0	3274	6	AX548756	AX548756 Sequence
42	c	21	100.0	6904	6	AX344977	AX344977 Sequence
43	c	21	100.0	6904	6	AX348498	AX348498 Sequence
44		21	100.0	9842	9	AF005900	AF005900 Homo sapi
45	c	21	100.0	22842	9	AC092603	AC092603 Homo sapi

ALIGNMENTS

RESULT 1
AX350501
LOCUS AX350501 21 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 13 from Patent WO0179561.
ACCESSION AX350501
VERSION AX350501.1 GI:18616097
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liggett,S.B. and Small,K.M.
TITLE Alpha-2 adrenergic receptor polymorphisms
JOURNAL Patent: WO 0179561-A 13 25-OCT-2001;

```
          Liggett, Stephen B. (US) ; Small, Kersten M. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT        2 a      9 c      2 g      8 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 6;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 30;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db          1 GCTCATCATCCCTTTCTCGCT 21


RESULT 2
AF337539
LOCUS       AF337539               1149 bp    DNA     linear   MAM 22-MAY-2001
DEFINITION  Pteropus rayneri alpha adrenergic receptor 2B (aar2B) gene, partial
            cds.
ACCESSION   AF337539
VERSION     AF337539.1  GI:14164929
KEYWORDS    .
SOURCE      Pteropus rayneri (Solomons flying fox)
  ORGANISM  Pteropus rayneri
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
            Pteropodinae; Pteropus.
REFERENCE   1  (bases 1 to 1149)
  AUTHORS   Springer,M.S., Teeling,E.C., Madsen,O., Stanhope,M.J. and de
            Jong,W.W.
  TITLE     Integrated fossil and molecular data reconstruct bat echolocation
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (11), 6241-6246 (2001)
  MEDLINE   21267399
   PUBMED   11353869
REFERENCE   2  (bases 1 to 1149)
  AUTHORS   Springer,M.S., Teeling,E.C., Madsen,O., Stanhope,M.J. and de
            Jong,W.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-2001) Department of Biology, University of
            California, 1354 Speith, Riverside, CA 92521, USA
FEATURES             Location/Qualifiers
     source          1. .1149
                     /organism="Pteropus rayneri"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:110942"
     gene            <1. .>1149
                     /gene="aar2B"
     mRNA            <1. .>1149
                     /gene="aar2B"
                     /product="alpha adrenergic receptor 2B"
     CDS             <1. .>1149
                     /gene="aar2B"
                     /note="A2AB"
                     /codon_start=1
                     /product="alpha adrenergic receptor 2B"
                     /protein_id="AAK55304.1"
                     /db_xref="GI:14164930"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLVYKGDQGPQPRGRPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPRAKGGLRDNESKQPHRVPGGPSTM
                     ASCLAASGEASRHSKPTGEKEQGETEDPXSXALPPSWPALPHAGQSPKEAVCGVSLEE
                     EXGEEEDECEPQALPASPASACSPPLQQPQGSRVLATLRGQVLLGRGMGTAGGQWWRR
                     RAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      183 a    386 c    340 g    235 t      5 others
ORIGIN

  Query Match           100.0%;  Score 21;  DB 4;  Length 1149;
  Best Local Similarity  100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        150 GCTCATCATCCCTTTCTCGCT 170


RESULT 3
RAM315937
LOCUS       RAM315937              1149 bp    DNA     linear   MAM 22-MAY-2002
DEFINITION  Rousettus amplexicaudatus partial adra2b gene for alpha 2B
            adrenergic receptor.
ACCESSION   AJ315937
VERSION     AJ315937.1  GI:21212918
KEYWORDS    adra2b gene; alpha 2B adrenergic receptor.
SOURCE      Rousettus amplexicaudatus (Geoffroy's rousette)
  ORGANISM  Rousettus amplexicaudatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
            Pteropodinae; Rousettus.
REFERENCE   1
  AUTHORS   Murphy,W.J., Eizirik,E., O'Brien,S.J., Madsen,O., Scally,M.,
            Douady,C.J., Teeling,E., Ryder,O.A., Stanhope,M.J., de Jong,W.W.
            and Springer,M.S.
  TITLE     Resolution of the early placental mammal radiation using Bayesian
            phylogenetics
  JOURNAL   Science 294 (5550), 2348-2351 (2001)
  MEDLINE   21608557
   PUBMED   11743200
REFERENCE   2  (bases 1 to 1149)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2001) Madsen O., 161 Biochemistry NWI, University
            of Nijmegen, P.O.Box 9101, 6500 HB Nijmegen, NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1149
                     /organism="Rousettus amplexicaudatus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:58083"
     gene            1. .1149
                     /gene="adra2b"
     CDS             <1. .>1149
                     /gene="adra2b"
                     /codon_start=1
                     /product="alpha 2B adrenergic receptor"
                     /protein_id="CAC87001.1"
                     /db_xref="GI:21212919"
                     /translation="AIAAAITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLVYKGDQGPQPHGRPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPRAKGGLRDSESKQPHRVPGGPSTL
                     ASCLAASGEASRHSKPTGEKEQGETEDPGSPALPPSWPALPHAGQNLKEAVCGVSLEE
                     NVEEEEDKCEPQALPASPASACSPPLQQPQGSRVLATLRGQVLLARGMGTAGGQWWRR
                     RAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      184 a    386 c    340 g    239 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 4;  Length 1149;
  Best Local Similarity  100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        150 GCTCATCATCCCTTTCTCGCT 170


RESULT 4
CSP251181
LOCUS       CSP251181              1150 bp    DNA     linear   MAM 01-JUN-2001
DEFINITION  Cynopterus sphinx partial aar2B gene for alpha adrenergic receptor
            2B.
ACCESSION   AJ251181
VERSION     AJ251181.1  GI:11322256
```

```
KEYWORDS    aar2B gene; alpha adrenergic receptor 2B.
SOURCE      Cynopterus sphinx (Indian short-nosed fruit bat)
  ORGANISM  Cynopterus sphinx
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
            Pteropodinae; Cynopterus.
REFERENCE   1
  AUTHORS   Madsen,O., Scally,M., Douady,C.J., Kao,D.J., DeBry,R.W., Adkins,R.,
            Amrine,H.M., Stanhope,M.J., de Jong,W.W. and Springer,M.S.
  TITLE     Parallel adaptive radiations in two major clades of placental
            mammals
  JOURNAL   Nature 409 (6820), 610-614 (2001)
  MEDLINE   21082081
   PUBMED   11214318
REFERENCE   2  (bases 1 to 1150)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1999) Madsen O., Department of Biochemistry,
            University of Nijmegen, P.O. box 9101, 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1150
                     /organism="Cynopterus sphinx"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9400"
     gene            1. .1150
                     /gene="aar2B"
     CDS             <1. .>1150
                     /gene="aar2B"
                     /codon_start=2
                     /product="alpha adrenergic receptor 2B"
                     /protein_id="CAC16684.1"
                     /db_xref="GI:11322257"
                     /db_xref="SPTREMBL:Q9GL20"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPHGAPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPRAKGGLRDSESKQPHRVPGGPSTL
                     ASCLPASGEAGRPSKPTGEKEQGETEDPGSPALPPSWPALPHGGQSPKEAVCGVSLEE
                     EIGEEEDECEPQALPASPASACSPPRQQPQGSRVLATLRGQVLLGKGLGTASGQWWRR
                     RAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      181 a    395 c    337 g    237 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 4;  Length 1150;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        151 GCTCATCATCCCTTTCTCGCT 171


RESULT 5
MLA251107
LOCUS       MLA251107               1156 bp    DNA     linear   MAM 15-NOV-2000
DEFINITION  Micropotamogale lamottei partial aar2B gene for alpha adrenergic
            receptor.
ACCESSION   AJ251107
VERSION     AJ251107.1  GI:11191767
KEYWORDS    aar2B gene; alpha adrenergic receptor; subtype 2B.
SOURCE      Micropotamogale lamottei (Nimba otter shrew)
  ORGANISM  Micropotamogale lamottei
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Insectivora; Tenrecidae; Potamogalinae;
            Micropotamogale.
REFERENCE   1
  AUTHORS   van Dijk,M.A.M., Madsen,O., Catzeflis,F., Stanhope,M. and de
            Jong,W.W.
  TITLE     Probing 1the protein sequence support for the 'African clade' of
            mammals
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1156)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-1999) Madsen O., Department of Biochemistry,
            University of Nijmegen, P.O. box 9101, 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1156
                     /organism="Micropotamogale lamottei"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:105689"
     gene            1. .1156
                     /gene="aar2B"
     CDS             <1. .>1156
                     /gene="aar2B"
                     /codon_start=2
                     /product="alpha adrenergic receptor, subtype 2B"
                     /protein_id="CAC16396.1"
                     /db_xref="GI:11191768"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRHTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCHILTVWLIAAAISLPPLIYKGDQGPQPHGRPQCKLNQEAWYI
                     LSSSIGSFFAPCLIMILVYLRIYLIAKRSSRCRGPGSKGTSRGGVGSKQPHPLAAGTS
                     AKPSTLTSSLAVAGEANGQSKPTGEEGETPKDPVAPTLPPNWAAHPHSGLGQKEGGTS
                     PEEEAEEEEECGPQAVPASPASAGSPPLQQPRGSRVLATLRGQVLLSRGLGAAGGQWW
                     RRRTQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      176 a    411 c    339 g    230 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 4;  Length 1156;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        151 GCTCATCATCCCTTTCTCGCT 171


RESULT 6
FCA251174
LOCUS       FCA251174               1162 bp    DNA     linear   MAM 01-JUN-2001
DEFINITION  Felis catus partial aar2B gene for alpha adrenergic receptor 2B.
ACCESSION   AJ251174
VERSION     AJ251174.1  GI:11322248
KEYWORDS    aar2B gene; alpha adrenergic receptor 2B.
SOURCE      Felis catus (cat)
  ORGANISM  Felis catus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE   1
  AUTHORS   Madsen,O., Scally,M., Douady,C.J., Kao,D.J., DeBry,R.W., Adkins,R.,
            Amrine,H.M., Stanhope,M.J., de Jong,W.W. and Springer,M.S.
  TITLE     Parallel adaptive radiations in two major clades of placental
            mammals
  JOURNAL   Nature 409 (6820), 610-614 (2001)
  MEDLINE   21082081
   PUBMED   11214318
REFERENCE   2  (bases 1 to 1162)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1999) Madsen O., Department of Biochemistry,
            University of Nijmegen, P.O. box 9101, 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1162
                     /organism="Felis catus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9685"
     gene            1. .1162
                     /gene="aar2B"
     CDS             <1. .>1162
                     /gene="aar2B"
                     /codon_start=2
                     /product="alpha adrenergic receptor 2B"
```

```
                     /protein_id="CAC16689.1"
                     /db_xref="GI:11322249"
                     /db_xref="SPTREMBL:Q9GL17"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPRGRPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPRAKGRPREGESKQPRPVPTGTSTK
                     MPTLASLAASGEANGHSKPTGAKEEGETPEDPATPALPPSWSALPNSGQGRKEGVCGA
                     SPEEEAEEEEEECEPRALPASPASACSPPLQQPQGSRVLATLRGQVLLGRGVGTSSGQ
                     WWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT       185 a    394 c    346 g    237 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 4;  Length 1162;
  Best Local Similarity 100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        151 GCTCATCATCCCTTTCTCGCT 171


RESULT 7
AF337538
LOCUS       AF337538                1167 bp    DNA     linear   MAM 22-MAY-2001
DEFINITION  Hipposideros commersoni alpha adrenergic receptor 2B (aar2B) gene,
            partial cds.
ACCESSION   AF337538
VERSION     AF337538.1  GI:14164927
KEYWORDS    .
SOURCE      Hipposideros commersoni (Commerson's roundleaf bat)
  ORGANISM  Hipposideros commersoni
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chiroptera; Microchiroptera; Rhinolophidae;
            Hipposiderinae; Hipposideros.
REFERENCE   1  (bases 1 to 1167)
  AUTHORS   Springer,M.S., Teeling,E.C., Madsen,O., Stanhope,M.J. and de
            Jong,W.W.
  TITLE     Integrated fossil and molecular data reconstruct bat echolocation
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (11), 6241-6246 (2001)
  MEDLINE   21267399
   PUBMED   11353869
REFERENCE   2  (bases 1 to 1167)
  AUTHORS   Springer,M.S., Teeling,E.C., Madsen,O., Stanhope,M.J. and de
            Jong,W.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-2001) Department of Biology, University of
            California, 1354 Speith, Riverside, CA 92521, USA
FEATURES             Location/Qualifiers
     source          1. .1167
                     /organism="Hipposideros commersoni"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:110941"
     gene            <1. .>1167
                     /gene="aar2B"
     mRNA            <1. .>1167
                     /gene="aar2B"
     CDS             /product="alpha adrenergic receptor 2B"
                     <1. .>1167
                     /gene="aar2B"
                     /note="A2AB"
                     /codon_start=1
                     /product="alpha adrenergic receptor 2B"
                     /protein_id="AAK55303.1"
                     /db_xref="GI:14164928"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLVYKDDPGPQPRGRPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPRAKGGPGEGESKQPHRVPRGASAK
                     LSTLASHQAASGEANGHTKPNGKKEQGETPEDPGSPTLPPTWPALPHAGQGPKEGVCG
                     VSPEEEAGEEEEDECEPQVLPASPASACRPPLQQPQGSQVLATLRGQVLLGRGMGAAG
                     GQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT       191 a    396 c    344 g    235 t      1 others
ORIGIN

  Query Match           100.0%;  Score 21;  DB 4;  Length 1167;
  Best Local Similarity 100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        150 GCTCATCATCCCTTTCTCGCT 170


RESULT 8
NAL419805
LOCUS       NAL419805               1167 bp    DNA     linear   MAM 08-FEB-2002
DEFINITION  Nyctimene albiventer partial adra2b gene for alpha 2B adrenergic
            receptor.
ACCESSION   AJ419805
VERSION     AJ419805.1  GI:18643971
KEYWORDS    adra2b gene; alpha 2B adrenergic receptor.
SOURCE      Nyctimene albiventer (common tube-nosed fruit bat)
  ORGANISM  Nyctimene albiventer
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chiroptera; Megachiroptera; Pteropodidae;
            Pteropodinae; Nyctimene.
REFERENCE   1
  AUTHORS   Teeling,E.C., Madsen,O., Van den Bussche,R.A., de Jong,W.W.,
            Stanhope,M.J. and Springer,M.S.
  TITLE     Microbat paraphyly and the convergent evolution of a key innovation
            in Old World rhinolophoid microbats
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1431-1436 (2002)
  MEDLINE   21819367
   PUBMED   11805285
REFERENCE   2  (bases 1 to 1167)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-2001) Madsen O., 161 Biochemistry, University of
            Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1167
                     /organism="Nyctimene albiventer"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:48988"
     gene            1. .1167
                     /gene="adra2b"
     CDS             <1. .>1167
                     /gene="adra2b"
                     /codon_start=1
                     /product="alpha 2B adrenergic receptor"
                     /protein_id="CAD11972.1"
                     /db_xref="GI:18643972"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPRGRPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPRAKGGLRDSESKQPHRVPGGPSTL
                     ASCLATSGEASRRSKPTKEKEQGETEDPGSPVLPPSWPDLPHAGQSLKEAVCGVSLEE
                     EVGEEEVGEEEDEGEPQDALSASPASACSPPLQQPQGSRVLATLRGQVLLGRGMGTAS
                     GQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT       192 a    378 c    348 g    249 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 4;  Length 1167;
  Best Local Similarity 100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        150 GCTCATCATCCCTTTCTCGCT 170


RESULT 9
RCR419806
LOCUS       RCR419806               1167 bp    DNA     linear   MAM 08-FEB-2002
DEFINITION  Rhinolophus creaghi partial adra2b gene for alpha 2B adrenergic
```

```
            receptor.
ACCESSION   AJ419806
VERSION     AJ419806.1  GI:18643973
KEYWORDS    adra2b gene; alpha 2B adrenergic receptor.
SOURCE      Rhinolophus creaghi
  ORGANISM  Rhinolophus creaghi
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chiroptera; Microchiroptera; Rhinolophidae;
            Rhinolophinae; Rhinolophus.
REFERENCE   1
  AUTHORS   Teeling,E.C., Madsen,O., Van den Bussche,R.A., de Jong,W.W.,
            Stanhope,M.J. and Springer,M.S.
  TITLE     Microbat paraphyly and the convergent evolution of a key innovation
            in Old World rhinolophoid microbats
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1431-1436 (2002)
  MEDLINE   21819367
   PUBMED   11805285
REFERENCE   2  (bases 1 to 1167)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-NOV-2001) Madsen O., 161 Biochemistry, University of
            Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1167
                     /organism="Rhinolophus creaghi"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:178895"
     gene            1. .1167
                     /gene="adra2b"
     CDS             <1. .>1167
                     /gene="adra2b"
                     /codon_start=1
                     /product="alpha 2B adrenergic receptor"
                     /protein_id="CAD11973.1"
                     /db_xref="GI:18643974"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCVILTVWLIAAVISLPPLVYKDDPGPQPRGRPQCKLNQEAWYI
                     LASSIASFFAPCLIMVLVYLRIYLIAKRSHRRGPRAKXGAGKDESKQPCRVARGASAK
                     LSTLASHEAASGEDNEHSKPNGEKEQGETPEDPGIPTLPPIWPALPHAGQGPTEGVCG
                     ASPEEDAGEEEEDECEPQVLRVSPASACSPPLQQPQGSRVLATLRGQVLLGRGMGAAG
                     GQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      195 a    391 c    344 g    235 t      2 others
ORIGIN

  Query Match           100.0%;  Score 21;  DB 4;  Length 1167;
  Best Local Similarity  100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        150 GCTCATCATCCCTTTCTCGCT 170


RESULT 10
PVI251176
LOCUS       PVI251176               1168 bp    DNA     linear   MAM 01-JUN-2001
DEFINITION  Phoca vitulina partial aar2B gene for alpha adrenergic receptor 2B.
ACCESSION   AJ251176
VERSION     AJ251176.1  GI:11322419
KEYWORDS    aar2B gene; alpha adrenergic receptor 2B.
SOURCE      Phoca vitulina (harbor seal)
  ORGANISM  Phoca vitulina
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Phoca.
REFERENCE   1
  AUTHORS   Madsen,O., Scally,M., Douady,C.J., Kao,D.J., DeBry,R.W., Adkins,R.,
            Amrine,H.M., Stanhope,M.J., de Jong,W.W. and Springer,M.S.
  TITLE     Parallel adaptive radiations in two major clades of placental
            mammals
  JOURNAL   Nature 409 (6820), 610-614 (2001)
  MEDLINE   21082081
   PUBMED   11214318
REFERENCE   2  (bases 1 to 1168)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1999) Madsen O., Department of Biochemistry,
            University of Nijmegen, P.O. box 9101, 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Phoca vitulina"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9720"
     gene            1. .1168
                     /gene="aar2B"
     CDS             <1. .>1168
                     /gene="aar2B"
                     /codon_start=2
                     /product="alpha adrenergic receptor 2B"
                     /protein_id="CAC16696.1"
                     /db_xref="GI:11322420"
                     /db_xref="SPTREMBL:Q9GL07"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPAGAPQCKLNQEAWYI
                     LASSIGSPFAPCLIMILVYLRIYLIAKRSHRRGPGAKGGPGPGESKQPRSVPTGTSTK
                     LPTLASLAASEEANGHSKPTGEKEGDTPEDPGTPALPPSWSALPSSGQSQKEGVCGAS
                     PEEEAEEEEEEEEECEPQALPASPASACSPPLPQPQGSRVLATLRGQVLLGRGVGTAS
                     GQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      189 a    398 c    351 g    230 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 4;  Length 1168;
  Best Local Similarity  100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        151 GCTCATCATCCCTTTCTCGCT 171


RESULT 11
AF337543
LOCUS       AF337543                1170 bp    DNA     linear   MAM 22-MAY-2001
DEFINITION  Taphozous sp. alpha adrenergic receptor 2B (aar2B) gene, partial
            cds.
ACCESSION   AF337543
VERSION     AF337543.1  GI:14164937
KEYWORDS    .
SOURCE      Taphozous sp.
  ORGANISM  Taphozous sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Chiroptera; Microchiroptera; Emballonuridae;
            Taphozous.
REFERENCE   1  (bases 1 to 1170)
  AUTHORS   Springer,M.S., Teeling,E.C., Madsen,O., Stanhope,M.J. and de
            Jong,W.W.
  TITLE     Integrated fossil and molecular data reconstruct bat echolocation
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (11), 6241-6246 (2001)
  MEDLINE   21267399
   PUBMED   11353869
REFERENCE   2  (bases 1 to 1170)
  AUTHORS   Springer,M.S., Teeling,E.C., Madsen,O., Stanhope,M.J. and de
            Jong,W.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-2001) Department of Biology, University of
            California, 1354 Speith, Riverside, CA 92521, USA
FEATURES             Location/Qualifiers
     source          1. .1170
                     /organism="Taphozous sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:110943"
     gene            <1. .>1170
                     /gene="aar2B"
     mRNA            <1. .>1170
```

```
                     /gene="aar2B"
                     /product="alpha adrenergic receptor 2B"
     CDS             <1. .>1170
                     /gene="aar2B"
                     /note="A2AB"
                     /codon_start=1
                     /product="alpha adrenergic receptor 2B"
                     /protein_id="AAK55308.1"
                     /db_xref="GI:14164938"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLVYKGDPGPQPRGRPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPRAKGSPGEGESKQPCRVPGRASTK
                     LPTLASHLAASGEANGHSKPTGEKEQGGTPEDPGSPALPPCWPALPHASQGPKEGVCG
                     ASPEEEVGEEEEDDECEPQALPASPASACSPPLQQPQGSRVLATLRGQVLLGRGMGAA
                     SGQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      194 a    388 c    347 g    241 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 4;  Length 1170;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        150 GCTCATCATCCCTTTCTCGCT 170


RESULT 12
DDAAR2B
LOCUS       DDAAR2B                 1171 bp    DNA     linear   MAM 15-NOV-1999
DEFINITION  Dugong dugon gene encoding alpha adrenergic receptor subtype 2B,
            partial.
ACCESSION   Y15947
VERSION     Y15947.2  GI:6448523
KEYWORDS    alpha adrenergic receptor; subtype 2B.
SOURCE      Dugong dugon (dugong)
  ORGANISM  Dugong dugon
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Sirenia; Dugongidae; Dugong.
REFERENCE   1
  AUTHORS   Stanhope,M.J., Madsen,O., Waddell,V.G., Cleven,G.C., de Jong,W.W.
            and Springer,M.S.
  TITLE     Highly congruent molecular support for a diverse superordinal clade
            of edemic African mammals
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-DEC-1997) O. Madsen, University of Nijmegen,
            Department of Biochemistry, PO Box 9101, NL-6500 HB Nijmegen,
            NETHERLANDS
  REMARK    Revised by [3]
REFERENCE   3  (bases 1 to 1171)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-NOV-1999) O. Madsen, University of Nijmegen,
            Department of Biochemistry, PO Box 9101, NL-6500 HB Nijmegen,
            NETHERLANDS
COMMENT     On Nov 17, 1999 this sequence version replaced gi:3288142.
FEATURES             Location/Qualifiers
     source          1. .1171
                     /organism="Dugong dugon"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:29137"
     CDS             <1. .>1171
                     /codon_start=2
                     /product="alpha adrenergic receptor subtype 2B"
                     /protein_id="CAA75900.2"
                     /db_xref="GI:6448524"
                     /db_xref="SWISS-PROT:O77713"
                     /translation="AIATVITFLILFTIFGNSLVILAVLTSRSLRAPQNLFLVSLAAA
                     DIMVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAAISLPPLIYKGDQGPQPRGRPQCKLNQEAWYI
                     LSSSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPGAKGGPRKGESKQPHSLDSGPSAL
                     ANLPTLASSLAVAGEANGHSMPPGEKERETSEDPGTPTLPPSWPVLPNSGQGQKGGVC
                     GASLEEEADKEEEEECGPPAVPASPATACNPPLQQPQGSQVLATLRGQVFLGRGVGAA
                     GGQWWRRWAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      189 a    397 c    337 g    247 t      1 others
ORIGIN

  Query Match             100.0%;  Score 21;  DB 4;  Length 1171;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        151 GCTCATCATCCCTTTCTCGCT 171


RESULT 13
TMA251109
LOCUS       TMA251109               1171 bp    DNA     linear   MAM 15-NOV-2000
DEFINITION  Trichechus manatus partial aar2B gene for alpha adrenergic
            receptor.
ACCESSION   AJ251109
VERSION     AJ251109.1  GI:11191771
KEYWORDS    aar2B gene; alpha adrenergic receptor; subtype 2B.
SOURCE      Trichechus manatus (Caribbean manatee)
  ORGANISM  Trichechus manatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Sirenia; Trichechidae; Trichechus.
REFERENCE   1
  AUTHORS   van Dijk,M.A.M., Madsen,O., Catzeflis,F., Stanhope,M. and de
            Jong,W.W.
  TITLE     Probing 1the protein sequence support for the 'African clade' of
            mammals
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1171)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-1999) Madsen O., Department of Biochemistry,
            University of Nijmegen, P.O. box 9101, 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1171
                     /organism="Trichechus manatus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9778"
     gene            1. .1171
                     /gene="aar2B"
     CDS             <1. .>1171
                     /gene="aar2B"
                     /codon_start=2
                     /product="alpha adrenergic receptor, subtype 2B"
                     /protein_id="CAC16418.1"
                     /db_xref="GI:11191772"
                     /translation="AIATVITFLILFTIFGNSLVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFWRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKRIILTVWLIAAAISLPPLIYKGDQGPQPRGRPQCKLNQEAWYI
                     LSSSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPGAKGAPRKGESKQPHSLDSGPSAL
                     ANLPTLASSLAVAGEADGHSMPPGEKERETSEDPGTPTLPPSWPVLPNSGQGQKEGVC
                     GASLEEEADKEEEEECGPPAVPASPATACSPPLQQPQGSRVLATLRGQVFLGRGVGAA
                     GGQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCRVPHGLF"
BASE COUNT      183 a    404 c    342 g    242 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 4;  Length 1171;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        151 GCTCATCATCCCTTTCTCGCT 171
```

```
RESULT 14
CVA251182
LOCUS       CVA251182               1171 bp    DNA     linear   PRI 01-JUN-2001
DEFINITION  Cynocephalus variegatus partial aar2B gene for alpha adrenergic
            receptor 2B.
ACCESSION   AJ251182
VERSION     AJ251182.1  GI:11322258
KEYWORDS    aar2B gene; alpha adrenergic receptor 2B.
SOURCE      Cynocephalus variegatus (Malayan flying lemur)
  ORGANISM  Cynocephalus variegatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Dermoptera; Cynocephalidae; Cynocephalus.
REFERENCE   1
  AUTHORS   Madsen,O., Scally,M., Douady,C.J., Kao,D.J., DeBry,R.W., Adkins,R.,
            Amrine,H.M., Stanhope,M.J., de Jong,W.W. and Springer,M.S.
  TITLE     Parallel adaptive radiations in two major clades of placental
            mammals
  JOURNAL   Nature 409 (6820), 610-614 (2001)
  MEDLINE   21082081
   PUBMED   11214318
REFERENCE   2  (bases 1 to 1171)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1999) Madsen O., Department of Biochemistry,
            University of Nijmegen, P.O. box 9101, 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1171
                     /organism="Cynocephalus variegatus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9457"
     gene            1. .1171
                     /gene="aar2B"
     CDS             <1. .>1171
                     /gene="aar2B"
                     /codon_start=2
                     /product="alpha adrenergic receptor 2B"
                     /protein_id="CAC16685.1"
                     /db_xref="GI:11322259"
                     /db_xref="SPTREMBL:Q9GL19"
                     /translation="AIAAVITFLILFTIFGNVLVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVAALIIPFSLANELLGYWYFWRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPLGRPQCKLNEEAWYI
                     LASSIGSFFAPCLIMILVYLRIYMIAKRSNRRGPRARGGPGEGESKQRRPGPGGALAS
                     AKLPAQASPLASAGEANGQSKPTGEKEEGEISEDPGTQTLPPSWTALPNSGQGQKEGV
                     CGASPGEEAEEAEEECEPQAVPVSSASICSPPLQQPQGSRVLATLRGHVLLGRGVGAV
                     GGQWWRRRAQLSREKRFTFVLAVVIGAFVLCWFPFFFSYSLGAICPQRCKVPHGLF"
BASE COUNT      173 a    388 c    369 g    241 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 9;  Length 1171;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        151 GCTCATCATCCCTTTCTCGCT 171


RESULT 15
ARU427256
LOCUS       ARU427256               1172 bp    DNA     linear   ROD 25-OCT-2002
DEFINITION  Aplodontia rufa partial A2AB gene for alpha 2B adrenergic receptor,
            exon 1.
ACCESSION   AJ427256
VERSION     AJ427256.1  GI:21655547
KEYWORDS    A2AB gene; alpha 2B adrenergic receptor.
SOURCE      Aplodontia rufa (mountain beaver)
  ORGANISM  Aplodontia rufa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Aplodontidae;
            Aplodontia.
REFERENCE   1
  AUTHORS   Huchon,D., Madsen,O., Sibbald,M.J., Ament,K., Stanhope,M.J.,
            Catzeflis,F., de Jong,W.W. and Douzery,E.J.
  TITLE     Rodent phylogeny and a timescale for the evolution of Glires:
            evidence from an extensive taxon sampling using three nuclear genes
  JOURNAL   Mol. Biol. Evol. 19 (7), 1053-1065 (2002)
  MEDLINE   22077630
   PUBMED   12082125
REFERENCE   2  (bases 1 to 1172)
  AUTHORS   Douzery,E.J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JAN-2002) Douzery E.J.P., Institut des Sciences de
            l'Evolution, Lab. Paleonto., Paleobio., Phylogenie, Universite
            Montpellier II, Pl. E. Bataillon, 34 095 Montpellier Cedex 5,
            FRANCE
FEATURES             Location/Qualifiers
     source          1. .1172
                     /organism="Aplodontia rufa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:51342"
     gene            1. .1172
                     /gene="A2AB"
     CDS             <1. .>1172
                     /gene="A2AB"
                     /codon_start=1
                     /product="alpha 2B adrenergic receptor"
                     /protein_id="CAD20294.1"
                     /db_xref="GI:21655548"
                     /translation="AAITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAADIL
                     VATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVSRAL
                     EYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPHGRPQCKLNQEAWYILAS
                     SIGSFFAPCLIMILVYLRIYLIAKRSHRRGPGAKGGPGKVVSKQSRPVPGGTPASAKV
                     PTLASPLSSPGEANGHPKPTGEKEEGETLEEPGTRALPASWTALPEADQGQKTSFCGA
                     SPEEEGEEEEEEEECDQCDGQSVPASSASVCSPPLQKPQGSRVLATLRGPVLLGRGVG
                     AVSGQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQRCKVPHGL"
     exon            <1. .>1172
                     /gene="A2AB"
                     /number=1
BASE COUNT      202 a    386 c    348 g    236 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 10;  Length 1172;
  Best Local Similarity   100.0%;  Pred. No. 18;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        141 GCTCATCATCCCTTTCTCGCT 161


Search completed: February 12, 2004, 04:41:36
Job time : 618.146 secs
```

This Page Blank (uspto)

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 01:53:19 ; Search time 110.378 Seconds
(without alignments)
513.582 Million cell updates/sec

Title: US-09-692-077D-13
Perfect score: 21
Sequence: 1 gctcatcatccctttctcgct 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	100.0	21	23	AAI99907	Human alpha-2BAR g
2	21	100.0	1344	22	AAD04761	Human alpha2B-adre
3	21	100.0	1344	23	AAI99906	Human alpha-2BAR t
4	21	100.0	1344	24	AAD44388	Human alpha-2B-adr
5	21	100.0	1353	22	AAD04762	Human alpha2B-adre
6	21	100.0	1353	23	AAI99905	Human alpha-2BAR t
7	21	100.0	1353	24	AAD44389	Human alpha-2B-adr
8	21	100.0	2072	25	ACA56583	Human signalling p
9	21	100.0	3274	25	ABZ42624	Human alpha 2b-adr
c 10	21	100.0	6904	24	ABL32075	Human immune syste
c 11	21	100.0	6904	24	AAD28365	Human chemically t
12	19.4	92.4	1350	23	AAI99917	Human alpha-2AAR e
13	19.4	92.4	1350	23	AAI99918	Human alpha-2AAR v
c 14	19.4	92.4	1733	24	ABQ47500	Oligonucleotide fo
15	19.4	92.4	1733	24	ABQ47501	Oligonucleotide fo
16	19.4	92.4	2064	12	AAQ14151	Human alpha 2 beta
17	19.4	92.4	2064	18	AAT59499	Human alpha-2b adr
18	19.4	92.4	3604	25	ACA56582	Human signalling p
19	19.4	92.4	3653	25	ABZ42623	Human alpha 2a-adr
c 20	19.4	92.4	7353	24	ABL32073	Human immune syste
c 21	19.4	92.4	7353	24	AAD28363	Human chemically t
c 22	18.4	87.6	297	25	ABX83493	Corn ear-derived p
c 23	17.8	84.8	3840	23	ABL18139	Drosophila melanog
c 24	17.8	84.8	7720	23	ABL18138	Drosophila melanog
c 25	16.8	80.0	261	21	AAC08191	Human secreted pro
c 26	16.8	80.0	376	23	ABV13690	Human prostate exp
c 27	16.8	80.0	401	23	ABV04521	Human prostate exp
c 28	16.8	80.0	401	23	ABV34803	Human prostate exp
c 29	16.8	80.0	401	23	ABV43654	Human prostate exp
c 30	16.8	80.0	967	21	AAA76276	Maize glutathione-
31	16.8	80.0	7327	23	ABL08624	Drosophila melanog
32	16.8	80.0	7933	23	ABL08720	Drosophila melanog
33	16.4	78.1	954	22	AAS42448	Human cDNA encodin
34	16.4	78.1	954	24	ABN89118	Human GPCR4a nucle
35	16.4	78.1	954	24	ABN89119	Human GPCR4b nucle
36	16.4	78.1	954	24	ABK37734	DNA encoding G-cou
37	16.4	78.1	958	24	ABT05663	GPCR 12 protein en
38	16.4	78.1	958	24	ABQ88380	Human G protein co
39	16.4	78.1	958	24	ABN89127	Human GPCR8a nucle
40	16.4	78.1	958	24	ABN89128	Human GPCR8b nucle
41	16.4	78.1	1017	22	AAS42422	Human cDNA encodin
42	16.4	78.1	1017	24	ABZ43146	Human GPCR polynuc
43	16.4	78.1	1017	24	ABK37708	DNA encoding G-cou
44	16.4	78.1	1049	22	AAH32040	Human olfactory re
45	16.4	78.1	1408	25	ABX17873	cDNA encoding huma

ALIGNMENTS

RESULT 1
AAI99907
ID AAI99907 standard; DNA; 21 BP.
XX
AC AAI99907;
XX
DT 18-FEB-2002 (first entry)
XX
DE Human alpha-2BAR genotyping PCR primer SEQ ID NO 13.
XX
KW Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW polymorphic site; allelic variant; cardiovascular disease;
KW central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW phosphorylation; inositol phosphate; alpha-2BAR; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200179561-A2.
XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;

XX
DR WPI; 2001-611728/70.
XX
PT Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT determining whether an individual is at increased risk of developing a
PT disease associated with the corresponding receptor comprises detecting
PT a polymorphic site -
XX
PS Claim 10; Page 112; 163pp; English.
XX
CC The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC alpha2A or alpha2C or fragment or complement of; and
CC (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC positions 961-972 of (III). The method may be used for genotyping an
CC alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC whether an individual is at increased risk of developing a disease
CC associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC polymorphic site which correlate to disease selected from cardiovascular
CC disease, central nervous system disease and combinations of these. In
CC addition, the technique may be used to predict an individual's response
CC to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC these) by detecting the polymorphic site and correlating the site to a
CC predetermined response (where the response is correlated to adenylyl
CC cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC levels). The present sequence is that of a human alpha-2BAR PCR primer,
CC useful for the genotyping methods of the invention.
XX
SQ Sequence 21 BP; 2 A; 9 C; 2 G; 8 T; 0 other;

```
  Query Match             100.0%;  Score 21;  DB 23;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db          1 GCTCATCATCCCTTTCTCGCT 21
```

RESULT 2
AAD04761
ID AAD04761 standard; DNA; 1344 BP.
XX
AC AAD04761;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human alpha2B-adrenoceptor (alpha2B-AR) variant gene.
XX
KW Human; cardiant; gene therapy; alpha2B-adrenoceptor; alpha2B-AR;
KW glutamic acid repeat; intracellular loop; chromosome 2; catecholamine;
KW norepinephrine; epinephrine; therapy; vascular contraction; variant;
KW coronary artery; coronary heat disease; CHD; chronic angina pectoris;
KW acute myocardial infarction; AMI; Prinzmetal's variant; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1344
FT /*tag= a
FT /product= "Human alpha2B-adrenoceptor (alpha2B-AR)
FT variant protein"
XX
PN WO200129082-A1.
XX
PD 26-APR-2001.
XX
PF 20-OCT-2000; 2000WO-FI00913.
XX
PR 22-OCT-1999; 99US-0422985.
XX
PA (JUVA-) JUVANTIA PHARMA LTD OY.
XX
PI Snapir A, Heinonen P, Alhopuro P, Karvonen M, Koulu M, Pesonen U;
PI Scheinin M, Salonen JT, Tuomainen T, Lakka TA, Nyyssoenen K;
PI Salonen R, Kauhanen J, Valkonen V;
XX
DR WPI; 2001-300318/31.
DR P-PSDB; AAE00989.
XX
PT New DNA molecule encoding variant specific adrenoceptor protein with
PT deletion of specific amino acids located in the third intracellular
PT loop of the polypeptide, for treating vascular contraction of coronary
PT arteries -
XX
PS Claim 3; Page 24-26; 37pp; English.
XX
CC The present sequence is a gene encoding human alpha2B-adrenoceptor
CC (alpha2B-AR) variant protein. Alpha2B-AR has a glutamic acid repeat
CC element (amino acids 298-309) of 12 glutamates, in an acidic stretch of
CC 18 amino acids (amino acids 294-311), located in the third intracellular
CC loop of the receptor polypeptide. The variant is obtained by deletion of
CC three glutamates from the Glu repeat (amino acids 307-309). Alpha2B-AR
CC gene is located on chromosome 2. Alpha2-AR mediate many of the
CC physiological effects of the catecholamines, norepinephrine and
CC epinephrine. An antagonist of alpha2B-adrenoceptor is useful for treating
CC a mammal suffering from vascular contraction of coronary arteries and a
CC disease involving vascular contraction of coronary arteries which is
CC clinically expressed as coronary heat disease (CHD), unstable chronic
CC angina pectoris which is clinically expressed as Prinzmetal's variant
CC form or acute myocardial infarction (AMI). Alpha2B-AR gene is used in
CC gene therapy.
XX
SQ Sequence 1344 BP; 219 A; 459 C; 400 G; 266 T; 0 other;

```
  Query Match             100.0%;  Score 21;  DB 22;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 3.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209
```

RESULT 3
AAI99906
ID AAI99906 standard; DNA; 1344 BP.
XX
AC AAI99906;
XX
DT 18-FEB-2002 (first entry)
XX
DE Human alpha-2BAR third intracellular loop variant encoding DNA.
XX
KW Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW polymorphic site; allelic variant; cardiovascular disease;
KW central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW phosphorylation; inositol phosphate; alpha-2BAR; chromosome 2; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1344
FT /*tag= a
FT /product= "alpha-2BAR"
FT /note= "sequence is deleted for a 9 nucleotide
FT polymorphic site found at nucleotides 901-909
FT of the wildtype alpha-2BAR protein (AAI99905)"
XX
PN WO200179561-A2.

XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;
XX
DR WPI; 2001-611728/70.
DR P-PSDB; AAM52118.
XX
PT Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT determining whether an individual is at increased risk of developing a
PT disease associated with the corresponding receptor comprises detecting
PT a polymorphic site -
XX
PS Claim 5; Page 144-145; 163pp; English.
XX
CC The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC alpha2A or alpha2C or fragment or complement of; and
CC (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC positions 961-972 of (III). The method may be used for genotyping an
CC alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC whether an individual is at increased risk of developing a disease
CC associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC polymorphic site which correlate to disease selected from cardiovascular
CC disease, central nervous system disease and combinations of these. In
CC addition, the technique may be used to predict an individual's response
CC to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC these) by detecting the polymorphic site and correlating the site to a
CC predetermined response (where the response is correlated to adenylyl
CC cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC levels). The present sequence is that of the third intracellular loop of
CC the human alpha-2BAR variant, the sequence is deleted for a 9 nucleotide
CC polymorphic site found at nucleotides 901-909 of the wildtype gene
CC (AAI99905).
XX
SQ Sequence 1344 BP; 220 A; 458 C; 400 G; 266 T; 0 other;

Query Match 100.0%; Score 21; DB 23; Length 1344;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCATCATCCCTTTCTCGCT 21
||||||||||||||||||||
Db 189 GCTCATCATCCCTTTCTCGCT 209

RESULT 4
AAD44388
ID AAD44388 standard; DNA; 1344 BP.
XX
AC AAD44388;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human alpha-2B-adrenoceptor variant DNA.
XX
KW Human; hypertension; alpha-2B-adrenoceptor; AR; antihypertensive;
KW hypertension; hypotensive; variant; gene; ds.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1344
FT /*tag= a
FT /product= "Human alpha-2B-adrenoceptor variant
FT protein"
XX
PN WO200266617-A1.
XX
PD 29-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-FI00113.
XX
PR 20-FEB-2001; 2001FI-0000323.
XX
PA (JURI-) JURILAB LTD OY.
XX
PI Salonen J;
XX
DR WPI; 2002-667063/71.
DR P-PSDB; AAE26633.
XX
PT Detecting a risk of hypertension and targeting treatment in a subject
PT by determining the pattern of alleles encoding a variant
PT alpha-2-adrenoceptor -
XX
PS Disclosure; Page 24-26; 35pp; English.
XX
CC The invention relates to a method for detecting a risk of hypertension
CC by determining the pattern of alleles encoding a variant alpha-2B-
CC adrenoceptor (AR) protein. The methods and compositions of the invention
CC are useful for detecting risks and targeting treatment for hypertension.
CC The kit is also useful for selecting for clinical drug trials testing
CC the antihypertensive effect of compounds. The present sequence is human
CC alpha-2B-adrenoceptor variant DNA.
XX
SQ Sequence 1344 BP; 219 A; 459 C; 400 G; 266 T; 0 other;

Query Match 100.0%; Score 21; DB 24; Length 1344;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCATCATCCCTTTCTCGCT 21
||||||||||||||||||||
Db 189 GCTCATCATCCCTTTCTCGCT 209

RESULT 5
AAD04762
ID AAD04762 standard; DNA; 1353 BP.
XX
AC AAD04762;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human alpha2B-adrenoceptor (alpha2B-AR) gene.
XX
KW Human; cardiant; gene therapy; alpha2B-adrenoceptor; alpha2B-AR;
KW glutamic acid repeat; intracellular loop; chromosome 2; catecholamine;
KW norepinephrine; epinephrine; therapy; vascular contraction;
KW coronary artery; coronary heat disease; CHD; chronic angina pectoris;
KW acute myocardial infarction; AMI; Prinzmetal's variant; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1353
FT /*tag= a
FT /product= "Human alpha2B-adrenoceptor (alpha2B-AR)
FT protein"

XX
PN WO200129082-A1.
XX
PD 26-APR-2001.
XX
PF 20-OCT-2000; 2000WO-FI00913.
XX
PR 22-OCT-1999; 99US-0422985.
XX
PA (JUVA-) JUVANTIA PHARMA LTD OY.
XX
PI Snapir A, Heinonen P, Alhopuro P, Karvonen M, Koulu M, Pesonen U;
PI Scheinin M, Salonen JT, Tuomainen T, Lakka TA, Nyyssoenen K;
PI Salonen R, Kauhanen J, Valkonen V;
XX
DR WPI; 2001-300318/31.
DR P-PSDB; AAE00990.
XX
PT New DNA molecule encoding variant specific adrenoceptor protein with
PT deletion of specific amino acids located in the third intracellular
PT loop of the polypeptide, for treating vascular contraction of coronary
PT arteries -
XX
PS Disclosure; Page 27-29; 37pp; English.
XX
CC The present sequence is a gene encoding human alpha2B-adrenoceptor
CC (alpha2B-AR) protein. Alpha2B-AR has a glutamic acid repeat element
CC (amino acids 298-309) of 12 glutamates, in an acidic stretch of 18 amino
CC acids (amino acids 294-311), located in the third intracellular loop of
CC the receptor polypeptide. Alpha2B-AR gene is located on chromosome 2.
CC Alpha2-AR mediate many of the physiological effects of the
CC catecholamines, norepinephrine and epinephrine. An antagonist of
CC alpha2B-adrenoceptor is useful for treating a mammal suffering from
CC vascular contraction of coronary arteries and a disease involving
CC vascular contraction of coronary arteries which is clinically expressed
CC as coronary heat disease (CHD), unstable chronic angina pectoris which is
CC clinically expressed as Prinzmetal's variant form or acute myocardial
CC infarction (AMI). Alpha2B-AR gene is used in gene therapy.
XX
SQ Sequence 1353 BP; 223 A; 459 C; 405 G; 266 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 1353;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCATCATCCCTTTCTCGCT 21
||||||||||||||||||||
Db 189 GCTCATCATCCCTTTCTCGCT 209

RESULT 6
AAI99905
ID AAI99905 standard; DNA; 1353 BP.
XX
AC AAI99905;
XX
DT 18-FEB-2002 (first entry)
XX
DE Human alpha-2BAR third intracellular loop encoding DNA.
XX
KW Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW polymorphic site; allelic variant; cardiovascular disease;
KW central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW phosphorylation; inositol phosphate; alpha-2BAR;
KW GenBank Accession AF009500; chromosome 2; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1353
FT /*tag= a
FT /product= "alpha-2BAR"
FT /note= "sequence includes a 9 nucleotide polymorphic site
FT at nucleotides 901-909 absent in the alpha-2BAR
FT variant (AAI99906)"
XX
PN WO200179561-A2.
XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;
XX
DR WPI; 2001-611728/70.
DR P-PSDB; AAM52117.
XX
PT Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT determining whether an individual is at increased risk of developing a
PT disease associated with the corresponding receptor comprises detecting
PT a polymorphic site -
XX
PS Claim 4; Page 144; 163pp; English.
XX
CC The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC alpha2A or alpha2C or fragment or complement of; and
CC (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC positions 961-972 of (III). The method may be used for genotyping an
CC alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC whether an individual is at increased risk of developing a disease
CC associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC polymorphic site which correlate to disease selected from cardiovascular
CC disease, central nervous system disease and combinations of these. In
CC addition, the technique may be used to predict an individual's response
CC to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC these) by detecting the polymorphic site and correlating the site to a
CC predetermined response (where the response is correlated to adenylyl
CC cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC levels). The present sequence is that of the third intracellular loop of
CC the human alpha-2BAR (GenBank Accession AF009500), the sequence includes
CC a 9 nucleotide polymorphic site at nucleotides 901-909, absent in the
CC alpha-2BAR variant (AAI99906).
XX
SQ Sequence 1353 BP; 224 A; 458 C; 405 G; 266 T; 0 other;

Query Match 100.0%; Score 21; DB 23; Length 1353;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCATCATCCCTTTCTCGCT 21
||||||||||||||||||||
Db 189 GCTCATCATCCCTTTCTCGCT 209

RESULT 7
AAD44389
ID AAD44389 standard; DNA; 1353 BP.
XX
AC AAD44389;
XX
DT 13-DEC-2002 (first entry)
XX

```
DE   Human alpha-2B-adrenoceptor gene.
XX
KW   Human; hypertension; alpha-2B-adrenoceptor; AR; antihypertensive;
KW   hypertension; hypotensive; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1353
FT                   /*tag= a
FT                   /product= "Human alpha-2B-adrenoceptor protein"
XX
PN   WO200266617-A1.
XX
PD   29-AUG-2002.
XX
PF   13-FEB-2002; 2002WO-FI00113.
XX
PR   20-FEB-2001; 2001FI-0000323.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Salonen J;
XX
DR   WPI; 2002-667063/71.
DR   P-PSDB; AAE26634.
XX
PT   Detecting a risk of hypertension and targeting treatment in a subject
PT   by determining the pattern of alleles encoding a variant
PT   alpha-2-adrenoceptor  -
XX
PS   Disclosure; Page 27-29; 35pp; English.
XX
CC   The invention relates to a method for detecting a risk of hypertension
CC   by determining the pattern of alleles encoding a variant alpha-2B-
CC   adrenoceptor (AR) protein. The methods and compositions of the invention
CC   are useful for detecting risks and targeting treatment for hypertension.
CC   The kit is also useful for selecting for clinical drug trials testing
CC   the antihypertensive effect of compounds. The present sequence is human
CC   alpha-2B-adrenoceptor gene.
XX
SQ   Sequence 1353 BP; 223 A; 459 C; 405 G; 266 T; 0 other;

  Query Match             100.0%;  Score 21;  DB 24;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 3.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209


RESULT 8
ACA56583
ID   ACA56583 standard; cDNA; 2072 BP.
XX
AC   ACA56583;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Human signalling pathway polynucleotide probe SEQ ID NO 1181.
XX
KW   Human; probe; ss; array element; Parkinson's disease;
KW   signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW   immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS   Homo sapiens.
XX
PN   US6500938-B1.
XX
PD   31-DEC-2002.
XX
PF   30-JAN-1998;   98US-0016434.
XX
PR   30-JAN-1998;   98US-0016434.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Au-Young J,  Seilhamer JJ;
XX
DR   WPI; 2003-352189/33.
XX
PT   Combination of polynucleotide probes, useful as array elements in a
PT   microarray for monitoring the expression of a number of target
PT   polynucleotides  -
XX
PS   Claim 1; SEQ ID NO 1181; 65pp; English.
XX
CC   The invention relates to a combination which, comprises a number of
CC   polynucleotide probes comprising a sequence selected from one of the 1490
CC   sequences mentioned in the specification. The combination is useful as an
CC   array element in a microarray for monitoring the expression of a number
CC   of target polynucleotides. The microarray is particularly useful in the
CC   diagnosis and treatment of cancer and immunopathology and neuropathology.
CC   The microarray is useful in diagnostics and treatment regimens, drug
CC   discovery and development, toxicological and carcinogenicity studies,
CC   forensics and pharmacogenomics. The microarray is also useful for
CC   monitoring progression of diseases and for developing sophisticated
CC   profiles for the effects of currently available therapeutic drugs. The
CC   combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC   and genomic fragments and in research and diagnostic applications. The
CC   array can detect changes in expression in a large number of genes coding
CC   for different signaling pathway populations which can be used to diagnose
CC   various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC   immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC   and Parkinson's disease. The present sequence represents a polynucleotide
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=06500938B1.
XX
SQ   Sequence 2072 BP; 316 A; 705 C; 660 G; 391 T; 0 other;

  Query Match             100.0%;  Score 21;  DB 25;  Length 2072;
  Best Local Similarity   100.0%;  Pred. No. 3.4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        601 GCTCATCATCCCTTTCTCGCT 621


RESULT 9
ABZ42624
ID   ABZ42624 standard; DNA; 3274 BP.
XX
AC   ABZ42624;
XX
DT   04-MAR-2003  (first entry)
XX
DE   Human alpha 2b-adrenoceptor nucleotide SEQ ID NO:41.
XX
KW   G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW   G protein-coupled receptor modulator; antibody; immune-related disease;
KW   growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW   immunological-related cell proliferative disease; autoimmune disease;
KW   Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW   osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW   graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW   psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW   mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW   hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW   ulcer; gene; ds.
XX
OS   Homo sapiens.
XX
```

PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US50107.
XX
PR 19-DEC-2000; 2000US-257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
DR P-PSDB; ABP81780.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating
PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
PT cancer or autoimmune diseases -
XX
PS Disclosure; Fig 1; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising:
CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular
CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity
CC or avidity for a particular GPCR. (I) can be used as GPCR modulators and
CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting
CC an antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention.
XX
SQ Sequence 3274 BP; 587 A; 979 C; 967 G; 741 T; 0 other;

Query Match 100.0%; Score 21; DB 25; Length 3274;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCATCATCCCTTTCTCGCT 21
||||||||||||||||||||||
Db 189 GCTCATCATCCCTTTCTCGCT 209

RESULT 10
ABL32075/c
ID ABL32075 standard; DNA; 6904 BP.
XX
AC ABL32075;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 48.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 48; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 6904 BP; 1326 A; 314 C; 2030 G; 3233 T; 1 other;

Query Match 100.0%; Score 21; DB 24; Length 6904;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCATCATCCCTTTCTCGCT 21
||||||||||||||||||||||
Db 1716 GCTCATCATCCCTTTCTCGCT 1696

RESULT 11
AAD28365/c
ID AAD28365 standard; DNA; 6904 BP.
XX
AC AAD28365;
XX
DT 22-APR-2002 (first entry)
XX
DE Human chemically treated genomic DNA #6.
XX
KW Human; cytostatic; antidepressant; neuroleptic; nootropic; antiaddictive;
KW adrenergic alpha-1C-receptor; cytosine methylation; therapy; alcoholism;
KW behavioural disorder; neurological; psychiatric; cancer; schizophrenia;
KW Tourette's syndrome; smoking; human immunodeficiency virus dementia;
KW drug abuse; migraine; ds.
XX
OS Homo sapiens.
XX
PN WO200202809-A2.
XX
PD 10-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07540.
XX
PR 30-JUN-2000; 2000DE-1032529.

```
PR   01-SEP-2000; 2000DE-1043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-154759/20.
XX
PT   Novel nucleic acid useful for diagnosis and therapy of behavioral
PT   disorder, neurological disorder and cancer, comprises a sequence of a
PT   segment of chemically pretreated DNA of adrenergic alpha-1C-receptor
PT   gene  -
XX
PS   Claim 1; Page 44-48; 190pp; English.
XX
CC   The invention relates to nucleic acids comprising a segment of chemically
CC   pretreated DNA of adrenergic alpha-1C-receptor gene. The invention also
CC   relates to oligonucleotides or peptide nucleic acid (PNA) oligomers
CC   useful for detecting cytosine methylations. The pretreated DNA is useful
CC   for the diagnosis or therapy of behavioural disorders, neurological
CC   disorders and cancer, in particular major depressive disorder, Tourette's
CC   syndrome, schizophrenia, psychiatric and neurological disorders, smoking,
CC   drug abuse, alcoholism, personality traits, compulsive gambling, human
CC   immunodeficiency virus dementia, migraine, behaviours in schizophrenic
CC   and schizoaffective patients, and suicidal behaviour in patients with
CC   schizophrenia. The nucleic acid is useful for detecting the methylation
CC   state of all CpG dinucleotides and/or single nucleotide polymorphisms
CC   (SNPs). The present sequence is human chemically treated genomic DNA.
XX
SQ   Sequence 6904 BP; 1326 A; 314 C; 2030 G; 3233 T; 1 other;

  Query Match             100.0%;  Score 21;  DB 24;  Length 6904;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db       1716 GCTCATCATCCCTTTCTCGCT 1696


RESULT 12
AAI99917
ID   AAI99917 standard; DNA; 1350 BP.
XX
AC   AAI99917;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Human alpha-2AAR encoding DNA.
XX
KW   Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW   polymorphic site; allelic variant; cardiovascular disease;
KW   central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW   phosphorylation; inositol phosphate; alpha-2AAR;
KW   GenBank Accession AF281308; chromosome 10; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1353
PT                   /*tag=  a
FT                   /product= "alpha-2AAR"
XX
PN   WO200179561-A2.
XX
PD   25-OCT-2001.
XX
PF   17-APR-2001; 2001WO-US12575.
XX
PR   17-APR-2000; 2000US-0551744.
PR   10-AUG-2000; 2000US-0636259.
PR   19-OCT-2000; 2000US-0692077.
XX
PA   (LIGG/) LIGGETT S B.
PA   (SMAL/) SMALL K M.
XX
PI   Liggett SB,  Small KM;
XX
DR   WPI; 2001-611728/70.
DR   P-PSDB; AAM52122.
XX
PT   Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT   determining whether an individual is at increased risk of developing a
PT   disease associated with the corresponding receptor comprises detecting
PT   a polymorphic site -
XX
PS   Example 7; Page 151; 163pp; English.
XX
CC   The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC   receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC   (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC   alpha2A or alpha2C or fragment or complement of; and
CC   (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC   of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC   or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC   positions 961-972 of (III). The method may be used for genotyping an
CC   alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC   whether an individual is at increased risk of developing a disease
CC   associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC   polymorphic site which correlate to disease selected from cardiovascular
CC   disease, central nervous system disease and combinations of these.. In
CC   addition, the technique may be used to predict an individual's response
CC   to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC   norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC   combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC   rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC   these) by detecting the polymorphic site and correlating the site to a
CC   predetermined response (where the response is correlated to adenylyl
CC   cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC   levels). The present sequence is that of the human alpha-2AAR gene
CC   (GenBank Accession AF281308).
XX
SQ   Sequence 1350 BP; 199 A; 490 C; 441 G; 220 T; 0 other;

  Query Match             92.4%;  Score 19.4;  DB 23;  Length 1350;
  Best Local Similarity   95.2%;  Pred. No. 18;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| |||||||||||||||
Db        252 GCTCGTCATCCCTTTCTCGCT 272


RESULT 13
AAI99918
ID   AAI99918 standard; DNA; 1350 BP.
XX
AC   AAI99918;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Human alpha-2AAR variant encoding DNA.
XX
KW   Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW   polymorphic site; allelic variant; cardiovascular disease;
KW   central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW   phosphorylation; inositol phosphate; alpha-2AAR; chromosome 10; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1353
FT                   /*tag=  a
FT                   /product= "alpha-2AAR"
FT   allele          replace(753,C)
FT                   /*tag=  b
```

XX
PN WO200179561-A2.
XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;
XX
DR WPI; 2001-611728/70.
DR P-PSDB; AAM52123.
XX
PT Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT determining whether an individual is at increased risk of developing a
PT disease associated with the corresponding receptor comprises detecting
PT a polymorphic site -
XX
PS Disclosure; Page 152; 163pp; English.
XX
CC The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC alpha2A or alpha2C or fragment or complement of; and
CC (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC positions 961-972 of (III). The method may be used for genotyping an
CC alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC whether an individual is at increased risk of developing a disease
CC associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC polymorphic site which correlate to disease selected from cardiovascular
CC disease, central nervous system disease and combinations of these. In
CC addition, the technique may be used to predict an individual's response
CC to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC these) by detecting the polymorphic site and correlating the site to a
CC predetermined response (where the response is correlated to adenylyl
CC cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC levels). The present sequence is that of the human alpha-2AAR variant
CC gene.
XX
SQ Sequence 1350 BP; 199 A; 489 C; 442 G; 220 T; 0 other;

Query Match 92.4%; Score 19.4; DB 23; Length 1350;
Best Local Similarity 95.2%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| ||||||||||||||||
Db        252 GCTCGTCATCCCTTTCTCGCT 272
```

RESULT 14
ABQ47500/c
ID ABQ47500 standard; DNA; 1733 BP.
XX
AC ABQ47500;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 34091.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP10074.
XX
PR 01-SEP-2000; 2000DE-1043826.
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 1733 BP; 271 A; 211 C; 625 G; 626 T; 0 other;

Query Match 92.4%; Score 19.4; DB 24; Length 1733;
Best Local Similarity 95.2%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| ||||||||||||||||
Db       1222 GCTCGTCATCCCTTTCTCGCT 1202
```

RESULT 15
ABQ47501
ID ABQ47501 standard; DNA; 1733 BP.
XX
AC ABQ47501;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 34092.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX

OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP10074.
XX
PR 01-SEP-2000; 2000DE-1043826.
PR 05-SEP-2000; 2000DE-1044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 1733 BP; 626 A; 625 C; 211 G; 271 T; 0 other;

Query Match 92.4%; Score 19.4; DB 24; Length 1733;
Best Local Similarity 95.2%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| |||||||||||||||
Db        512 GCTCGTCATCCCTTTCTCGCT 532
```

Search completed: February 12, 2004, 04:01:23
Job time : 111.378 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 04:01:41 ; Search time 28.6829 Seconds
(without alignments)
323.156 Million cell updates/sec

Title: US-09-692-077D-13
Perfect score: 21
Sequence: 1 gctcatcatccctttctcgct 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21	100.0	2072	4	US-09-016-434-1181	Sequence 1181, Ap
	2	19.4	92.4	3604	4	US-09-016-434-1180	Sequence 1180, Ap
c	3	18.4	87.6	297	4	US-09-313-294A-1953	Sequence 1953, Ap
c	4	16.8	80.0	967	3	US-09-248-335-51	Sequence 51, Appl
c	5	16.8	80.0	1500	4	US-09-252-991A-16352	Sequence 16352, A
	6	16.8	80.0	1584	4	US-09-252-991A-16239	Sequence 16239, A
c	7	16.4	78.1	2102	1	US-08-784-651-5	Sequence 5, Appli
c	8	16.2	77.1	1860	4	US-09-489-847-53	Sequence 53, Appl
	9	16.2	77.1	2837	2	US-08-993-228-11	Sequence 11, Appl
c	10	16	76.2	3430	4	US-09-462-561B-9	Sequence 9, Appli
c	11	15.8	75.2	414	4	US-09-511-625B-21	Sequence 21, Appl
c	12	15.8	75.2	1322	4	US-09-634-238-178	Sequence 178, App
c	13	15.8	75.2	2168	1	US-08-784-651-1	Sequence 1, Appli
	14	15.8	75.2	2323	4	US-09-342-647-17	Sequence 17, Appl
c	15	15.8	75.2	17425	4	US-09-511-625B-5	Sequence 5, Appli
	16	15.4	73.3	8532	1	US-08-452-655B-1	Sequence 1, Appli
	17	15.4	73.3	8532	3	US-08-450-582-1	Sequence 1, Appli
	18	15.4	73.3	9606	1	US-07-741-940-1	Sequence 1, Appli
	19	15.4	73.3	9606	1	US-08-289-548A-1	Sequence 1, Appli
	20	15.4	73.3	9606	1	US-08-452-654-1	Sequence 1, Appli
	21	15.4	73.3	9606	2	US-08-370-235A-1	Sequence 1, Appli
	22	15.4	73.3	9606	4	US-08-449-731-1	Sequence 1, Appli
c	23	15.4	73.3	580073	4	US-08-545-528D-1	Sequence 1, Appli
	24	15.2	72.4	2573	4	US-09-620-312D-609	Sequence 609, App
c	25	15.2	72.4	4344	4	US-09-462-561B-11	Sequence 11, Appl
c	26	15.2	72.4	6545	5	PCT-US95-13749-3	Sequence 3, Appli
	27	15.2	72.4	8815	4	US-09-687-731-12	Sequence 12, Appl
c	28	15.2	72.4	36519	3	US-08-923-137-2	Sequence 2, Appli
	29	15.2	72.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	30	15.2	72.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	31	14.8	70.5	155	3	US-08-444-818-1	Sequence 1, Appli
c	32	14.8	70.5	155	3	US-08-444-818-3	Sequence 3, Appli
c	33	14.8	70.5	164	3	US-08-444-818-6	Sequence 6, Appli
c	34	14.8	70.5	318	3	US-08-444-818-5	Sequence 5, Appli
c	35	14.8	70.5	353	3	US-08-444-818-4	Sequence 4, Appli
c	36	14.8	70.5	353	3	US-08-444-818-9	Sequence 9, Appli
c	37	14.8	70.5	387	3	US-08-444-818-7	Sequence 7, Appli
c	38	14.8	70.5	617	3	US-09-146-221-9	Sequence 9, Appli
c	39	14.8	70.5	712	3	US-08-444-818-13	Sequence 13, Appl
c	40	14.8	70.5	943	2	US-08-483-695-6	Sequence 6, Appli
c	41	14.8	70.5	943	2	US-08-483-695-43	Sequence 43, Appl
c	42	14.8	70.5	943	2	US-07-965-285-6	Sequence 6, Appli
c	43	14.8	70.5	943	2	US-07-965-285-43	Sequence 43, Appl
c	44	14.8	70.5	943	2	US-08-487-231-6	Sequence 6, Appli
c	45	14.8	70.5	943	2	US-08-487-231-43	Sequence 43, Appl

ALIGNMENTS

```
RESULT 1
US-09-016-434-1181
; Sequence 1181, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1181:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2072 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g178197
US-09-016-434-1181

  Query Match             100.0%;  Score 21;  DB 4;  Length 2072;
  Best Local Similarity   100.0%;  Pred. No. 0.48;
```

```
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        601 GCTCATCATCCCTTTCTCGCT 621


RESULT 2
US-09-016-434-1180
; Sequence 1180, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1180:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3604 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g178195
US-09-016-434-1180

  Query Match             92.4%;  Score 19.4;  DB 4;  Length 3604;
  Best Local Similarity   95.2%;  Pred. No. 2.9;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| |||||||||||||||
Db       2329 GCTCGTCATCCCTTTCTCGCT 2349


RESULT 3
US-09-313-294A-1953/c
; Sequence 1953, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 1953
;    LENGTH: 297
;    TYPE: DNA
;    ORGANISM: Zea mays
;    FEATURE:
;    NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. 6476212 700551803H1
US-09-313-294A-1953

  Query Match             87.6%;  Score 18.4;  DB 4;  Length 297;
  Best Local Similarity   95.0%;  Pred. No. 6.1;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CTCATCATCCCTTTCTCGCT 21
              ||||||||||||||| ||||
Db        235 CTCATCATCCCTTTCACGCT 216


RESULT 4
US-09-248-335-51/c
; Sequence 51, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
;  APPLICANT: MCGONIGLE, BRIAN
;  APPLICANT: O'KEEF, DANIEL
;  TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
;  FILE REFERENCE: CL-1128-A
;  CURRENT APPLICATION NUMBER: US/09/248,335
;  CURRENT FILING DATE: 1999-02-10
;  EARLIER APPLICATION NUMBER: 08/924,759
;  EARLIER FILING DATE: 1997-September-05
;  NUMBER OF SEQ ID NOS: 74
;  SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 51
;    LENGTH: 967
;    TYPE: DNA
;    ORGANISM: maize
US-09-248-335-51

  Query Match             80.0%;  Score 16.8;  DB 3;  Length 967;
  Best Local Similarity   90.0%;  Pred. No. 40;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGC 20
              |||||||||| ||| |||||
Db         28 GCTCATCATCGCTTGCTCGC 9


RESULT 5
US-09-252-991A-16352/c
; Sequence 16352, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 16352
;    LENGTH: 1500
;    TYPE: DNA
;    ORGANISM: Pseudomonas aeruginosa
```

US-09-252-991A-16352

Query Match 80.0%; Score 16.8; DB 4; Length 1500;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GCTCATCATCCCTTTCTCGC 20
              |||||||| |||||| ||||
Db       1434 GCTCATCACCCCTTTGTCGC 1415
```

RESULT 6
US-09-252-991A-16239
; Sequence 16239, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16239
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16239

Query Match 80.0%; Score 16.8; DB 4; Length 1584;
Best Local Similarity 90.0%; Pred. No. 42;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GCTCATCATCCCTTTCTCGC 20
              |||||||| |||||| ||||
Db        124 GCTCATCACCCCTTTGTCGC 143
```

RESULT 7
US-08-784-651-5/c
; Sequence 5, Application US/08784651
; Patent No. 5821102
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy
; APPLICANT: Boominathan, Karuppan
; APPLICANT: Sandal, Thomas
; TITLE OF INVENTION: Nucleic Acids Encoding Polypeptides
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5821102o No. 5821102disk of No. 5821102th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,651
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Agris, Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 4608.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2102 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-784-651-5

Query Match 78.1%; Score 16.4; DB 1; Length 2102;
Best Local Similarity 94.4%; Pred. No. 68;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          3 TCATCATCCCTTTCTCGC 20
              |||||||||| |||||||
Db        336 TCATCATCCCGTTCTCGC 319
```

RESULT 8
US-09-489-847-53/c
; Sequence 53, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-847-53

Query Match 77.1%; Score 16.2; DB 4; Length 1860;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||  ||||| ||
Db         38 GCTCATCATCCTCTTCTCCCT 18
```

RESULT 9
US-08-993-228-11
; Sequence 11, Application US/08993228
; Patent No. 5976838
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

```
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/993,228
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sprunger, Suzanne A.
;      REGISTRATION NUMBER:  41,323
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8284
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2837 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-993-228-11

  Query Match             77.1%;  Score 16.2;  DB 2;  Length 2837;
  Best Local Similarity   85.7%;  Pred. No. 87;
  Matches   18;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||| | |||| |||
Db        625 GCTCATCATCGCCTTCTTGCT 645


RESULT 10
US-09-462-561B-9/c
; Sequence 9, Application US/09462561B
; Patent No. 6455252
; GENERAL INFORMATION:
;  APPLICANT: Wade, Nicholas M.
;  APPLICANT:  Harrison, Bruce T.
;  APPLICANT:  King, Brian W.
;  APPLICANT:  Reed, Kenneth C.
;  APPLICANT:  Murphy, Kathleen M.
;  TITLE OF INVENTION: DETERMINATION OF GENETIC SEX IN EQUINE SPECIES BY
;  TITLE OF INVENTION:  ANALYSIS OF Y-CHROMOSOMAL DNA SEQUENCES
;  FILE REFERENCE: Wade et al
;  CURRENT APPLICATION NUMBER: US/09/462,561B
;  CURRENT FILING DATE:  2000-03-22
;  PRIOR APPLICATION NUMBER: P07802
;  PRIOR FILING DATE: 1997-07-09
;  PRIOR APPLICATION NUMBER: PCT/AU98/00533
;  PRIOR FILING DATE: 1998-07-08
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;  LENGTH: 3430
;  TYPE: DNA
;  ORGANISM: Equus caballus
;  FEATURE:
;  NAME/KEY: unsure
;  LOCATION: (1110)
;  OTHER INFORMATION: n at position 1110 = a, t, c, or g
;  NAME/KEY: unsure
;  LOCATION: (1656)
;  OTHER INFORMATION: n at position 1656 = a, t, c, or g
;  NAME/KEY: unsure
;  LOCATION: (1658)
;  OTHER INFORMATION: n at position 1658 = a, t, c, or g
;  NAME/KEY: unsure
;  LOCATION: (1995)
;  OTHER INFORMATION: n at position 1995 = a, t, c, or g
;  NAME/KEY: unsure
;  LOCATION: (2087)
;  OTHER INFORMATION: n at position 2087 = a, t, c, or g
;  NAME/KEY: unsure
;  LOCATION: (2781)
;  OTHER INFORMATION: n at position 2781 = a, t, c, or g
US-09-462-561B-9

  Query Match             76.2%;  Score 16;  DB 4;  Length 3430;
  Best Local Similarity   94.1%;  Pred. No. 1.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CTCATCATCCCTTTCTC 18
              || |||||||||||||||
Db       2089 CTNATCATCCCTTTCTC 2073


RESULT 11
US-09-511-625B-21/c
; Sequence 21, Application US/09511625B
; Patent No. 6368828
; GENERAL INFORMATION:
;  APPLICANT:  LaRochelle, William J.
;  APPLICANT:  Patel, Bharvin
;  APPLICANT:  Pierce, Jacalyn H.
;  TITLE OF INVENTION:  ATTENUATED AND DOMINANT NEGATIVEVARIANT
;  TITLE OF INVENTION:  cDNAs OF STAT6: STAT6b AND STAT6c
;  FILE REFERENCE:  14014.0300u1
;  CURRENT APPLICATION NUMBER: US/09/511,625B
;  CURRENT FILING DATE:  2000-02-23
;  PRIOR APPLICATION NUMBER:  PCT/US98/17821
;  PRIOR FILING DATE:  1998-08-27
;  PRIOR APPLICATION NUMBER:  60/070,397
;  PRIOR FILING DATE:  1998-01-05
;  PRIOR APPLICATION NUMBER:  60/056,075
;  PRIOR FILING DATE:  1997-08-27
;  NUMBER OF SEQ ID NOS:  70
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 21
;  LENGTH: 414
;  TYPE: DNA
;  ORGANISM: Artificial Sequence
;  FEATURE:
;  OTHER INFORMATION: Description of artificial Sequence:/No. 6368828e -
;  OTHER INFORMATION: synthetic construct
US-09-511-625B-21

  Query Match             75.2%;  Score 15.8;  DB 4;  Length 414;
  Best Local Similarity   89.5%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 TCATCATCCCTTTCTCGCT 21
              ||| ||| ||||||||||||
Db        340 TCAGCATTCCTTTCTCGCT 322


RESULT 12
US-09-634-238-178/c
; Sequence 178, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
;  APPLICANT:  Glenn, Matthew
;  APPLICANT:  Havukkala, Ilkka J.
;  APPLICANT:  Bloksberg, Leonard, N.
;  APPLICANT:  Lubbers, Mark W.
;  APPLICANT:  Dekker, James
```

```
;  APPLICANT:  Christensson, Anna C.
;  APPLICANT:  Holland, Ross
;  APPLICANT:  O'Toole, Paul W.
;  APPLICANT:  Reid, Julian R.
;  APPLICANT:  Coolbear, Timothy
;  TITLE OF INVENTION:  Polynucleotides, materials incorporating
;  TITLE OF INVENTION:  them and methods for using them.
;  FILE REFERENCE:  11000.1043U1
;  CURRENT APPLICATION NUMBER: US/09/634,238
;  CURRENT FILING DATE:  2000-08-08
;  NUMBER OF SEQ ID NOS:  422
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 178
;   LENGTH: 1322
;   TYPE: DNA
;   ORGANISM: Lactobacillus rhamnosus
US-09-634-238-178

  Query Match             75.2%;  Score 15.8;  DB 4;  Length 1322;
  Best Local Similarity   89.5%;  Pred. No. 1.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 TCATCATCCCTTTCTCGCT 21
              ||||||| || ||||||||
Db        288 TCATCATACCGTTCTCGCT 270


RESULT 13
US-08-784-651-1/c
; Sequence 1, Application US/08784651
; Patent No. 5821102
;  GENERAL INFORMATION:
;    APPLICANT:  Berka, Randy
;    APPLICANT:  Boominathan, Karuppan
;    APPLICANT:  Sandal, Thomas
;    TITLE OF INVENTION:  Nucleic Acids Encoding Polypeptides
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5821102o No. 5821102disk of No. 5821102th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,651
;      FILING DATE:  21-JAN-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Agris, Cheryl H.
;      REGISTRATION NUMBER:  34,086
;      REFERENCE/DOCKET NUMBER:  4608.200-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2168 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-08-784-651-1

  Query Match             75.2%;  Score 15.8;  DB 1;  Length 2168;
  Best Local Similarity   89.5%;  Pred. No. 1.3e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CTCATCATCCCTTTCTCGC 20
              |||| |||||| |||||||
Db        346 CTCAACATCCCGTTCTCGC 328


RESULT 14
US-09-342-647-17
; Sequence 17, Application US/09342647A
; Patent No. 6368840
; GENERAL INFORMATION:
;  APPLICANT: Cahoon, Edgar B.
;  APPLICANT: Cahoon, Rebecca E.
;  APPLICANT: Hitz, William D.
;  APPLICANT: Kinney, Anthony J.
;  TITLE OF INVENTION: ACYL-COA Oxidase Homologs
;  FILE REFERENCE: BB-1175
;  CURRENT APPLICATION NUMBER: US/09/342,647A
;  CURRENT FILING DATE: 1999-06-29
;  EARLIER APPLICATION NUMBER: 60/092,482
;  EARLIER FILING DATE: 1998-07-10
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
;   LENGTH: 2323
;   TYPE: DNA
;   ORGANISM: Oryza sativa
US-09-342-647-17

  Query Match             75.2%;  Score 15.8;  DB 4;  Length 2323;
  Best Local Similarity   89.5%;  Pred. No. 1.3e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 TCATCATCCCTTTCTCGCT 21
              ||||| ||||||||| |||
Db       2064 TCATCCTCCCTTTCTTGCT 2082


RESULT 15
US-09-511-625B-5/c
; Sequence 5, Application US/09511625B
; Patent No. 6368828
; GENERAL INFORMATION:
;  APPLICANT:  LaRochelle, William J.
;  APPLICANT:  Patel, Bharvin
;  APPLICANT:  Pierce, Jacalyn H.
;  TITLE OF INVENTION:  ATTENUATED AND DOMINANT NEGATIVEVARIANT
;  TITLE OF INVENTION:  cDNAs OF STAT6: STAT6b AND STAT6c
;  FILE REFERENCE:  14014.0300u1
;  CURRENT APPLICATION NUMBER: US/09/511,625B
;  CURRENT FILING DATE:  2000-02-23
;  PRIOR APPLICATION NUMBER:  PCT/US98/17821
;  PRIOR FILING DATE:  1998-08-27
;  PRIOR APPLICATION NUMBER:  60/070,397
;  PRIOR FILING DATE:  1998-01-05
;  PRIOR APPLICATION NUMBER:  60/056,075
;  PRIOR FILING DATE:  1997-08-27
;  NUMBER OF SEQ ID NOS:  70
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 17425
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial Sequence:/No. 6368828e -
;   OTHER INFORMATION: synthetic construct
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(17425)
;   OTHER INFORMATION: n = a, t, c or g
US-09-511-625B-5

  Query Match             75.2%;  Score 15.8;  DB 4;  Length 17425;
  Best Local Similarity   89.5%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          3 TCATCATCCCTTTCTCGCT 21
              ||| ||| ||||||||||||
Db       6790 TCAGCATTCCTTTCTCGCT 6772
```

Search completed: February 12, 2004, 06:07:47
Job time : 40.6829 secs

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 04:42:03 ; Search time 119.341 Seconds
(without alignments)
648.193 Million cell updates/sec

Title: US-09-692-077D-13
Perfect score: 21
Sequence: 1 gctcatcatccctttctcgct 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	100.0	21	15	US-10-001-073-13	Sequence 13, Appl
2	21	100.0	1344	9	US-09-825-923-1	Sequence 1, Appli
3	21	100.0	1344	15	US-10-077-870-1	Sequence 1, Appli
4	21	100.0	1344	15	US-10-001-073-2	Sequence 2, Appli
5	21	100.0	1353	9	US-09-825-923-3	Sequence 3, Appli
6	21	100.0	1353	15	US-10-077-870-3	Sequence 3, Appli
7	21	100.0	1353	15	US-10-001-073-1	Sequence 1, Appli
8	21	100.0	2072	12	US-10-305-720-1181	Sequence 1181, Ap
9	21	100.0	3274	15	US-10-225-567A-41	Sequence 41, Appl
c 10	21	100.0	6904	13	US-10-311-455-48	Sequence 48, Appl
11	19.4	92.4	1350	15	US-10-001-073-24	Sequence 24, Appl
12	19.4	92.4	1350	15	US-10-001-073-25	Sequence 25, Appl
13	19.4	92.4	3604	12	US-10-305-720-1180	Sequence 1180, Ap
14	19.4	92.4	3653	15	US-10-225-567A-39	Sequence 39, Appl
c 15	19.4	92.4	7353	13	US-10-311-455-46	Sequence 46, Appl
16	17.4	82.9	96588	12	US-10-085-117-64	Sequence 64, Appl
17	17	81.0	688	13	US-10-027-632-173743	Sequence 173743,
18	17	81.0	688	14	US-10-027-632-173743	Sequence 173743,
19	16.8	80.0	628	13	US-10-027-632-224986	Sequence 224986,
20	16.8	80.0	628	14	US-10-027-632-224986	Sequence 224986,
21	16.4	78.1	954	10	US-09-886-055-480	Sequence 480, App
22	16.4	78.1	954	11	US-09-804-291-480	Sequence 480, App
23	16.4	78.1	954	13	US-09-965-422-13	Sequence 13, Appl
24	16.4	78.1	954	13	US-09-965-422-15	Sequence 15, Appl
25	16.4	78.1	958	12	US-10-005-041A-29	Sequence 29, Appl
26	16.4	78.1	958	13	US-10-025-806-53	Sequence 53, Appl
27	16.4	78.1	958	13	US-09-965-422-31	Sequence 31, Appl
28	16.4	78.1	958	13	US-09-965-422-33	Sequence 33, Appl
29	16.4	78.1	1017	10	US-09-886-055-428	Sequence 428, App
30	16.4	78.1	1017	11	US-09-804-291-428	Sequence 428, App
31	16.4	78.1	1354	12	US-10-292-798-1	Sequence 1, Appli
32	16.4	78.1	1354	13	US-10-017-161-1	Sequence 1, Appli
33	16.4	78.1	2631	10	US-09-801-368-43	Sequence 43, Appl
34	16.4	78.1	26928	13	US-10-374-979-6	Sequence 6, Appli
c 35	16.2	77.1	268	9	US-09-294-093B-2749	Sequence 2749, Ap
c 36	16.2	77.1	509	12	US-10-260-238-3223	Sequence 3223, Ap
37	16.2	77.1	540	10	US-09-974-300-6869	Sequence 6869, Ap
c 38	16.2	77.1	689	13	US-10-027-632-134859	Sequence 134859,
c 39	16.2	77.1	689	14	US-10-027-632-134859	Sequence 134859,
40	16.2	77.1	714	13	US-10-027-632-14649	Sequence 14649, A
41	16.2	77.1	714	14	US-10-027-632-14649	Sequence 14649, A
42	16.2	77.1	966	10	US-09-886-055-124	Sequence 124, App
43	16.2	77.1	966	11	US-09-804-291-124	Sequence 124, App
44	16.2	77.1	1121	13	US-09-844-861A-3	Sequence 3, Appli
45	16.2	77.1	1311	12	US-10-369-493-28507	Sequence 28507, A

ALIGNMENTS

RESULT 1
US-10-001-073-13
; Sequence 13, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
; APPLICANT: Liggett, Stephen
; APPLICANT: Small, Kirsten
; TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
; FILE REFERENCE: 13073-PCT
; CURRENT APPLICATION NUMBER: US/10/001,073
; CURRENT FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-001-073-13

Query Match 100.0%; Score 21; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db          1 GCTCATCATCCCTTTCTCGCT 21
```

RESULT 2
US-09-825-923-1
; Sequence 1, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
; APPLICANT: Snapir, Amir
; APPLICANT: Heinonen, Paula
; APPLICANT: Alhopuro, Pia
; APPLICANT: Karvonen, Matti
; APPLICANT: Koulu, Markku

```
;  APPLICANT:  Pesonen, Ullamari
;  APPLICANT:  Scheinin, Mika
;  APPLICANT:  Salonen, Jukka T
;  APPLICANT:  Tuomainen, Tomi-Pekka
;  APPLICANT:  Lakka, Timo A
;  APPLICANT:  Nyyss"nen, Kristiina
;  APPLICANT:  Salonen, Riitta
;  APPLICANT:  Kauhanen, Jussi
;  APPLICANT:  Valkonen, Veli-Pekka
;  TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
;  TITLE OF INVENTION:  protein, and uses thereof
;  FILE REFERENCE: Alpha-2B-AR variant
;  CURRENT APPLICATION NUMBER: US/09/825,923
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/422,985
;  PRIOR FILING DATE: 2000-05-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1341)
;   OTHER INFORMATION: Coding sequence for variant human
;   OTHER INFORMATION: alpha-2B-adrenoceptor protein
US-09-825-923-1

  Query Match             100.0%;  Score 21;  DB 9;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209


RESULT 3
US-10-077-870-1
; Sequence 1, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
;  APPLICANT: Salonen, Jukka T
;  TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
;  FILE REFERENCE: 0933-0183P
;  CURRENT APPLICATION NUMBER: US/10/077,870
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: FI 20010323
;  PRIOR FILING DATE: 2001-02-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1341)
;   OTHER INFORMATION: Coding sequence for variant human alpha-2B-adrenoceptor protein
US-10-077-870-1

  Query Match             100.0%;  Score 21;  DB 15;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209


RESULT 4
US-10-001-073-2
; Sequence 2, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
;  APPLICANT: Liggett, Stephen
;  APPLICANT:  Small, Kirsten
;  TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
;  FILE REFERENCE: 13073-PCT
;  CURRENT APPLICATION NUMBER: US/10/001,073
;  CURRENT FILING DATE:  2001-11-01
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-001-073-2

  Query Match             100.0%;  Score 21;  DB 15;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209


RESULT 5
US-09-825-923-3
; Sequence 3, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
;  APPLICANT: Snapir, Amir
;  APPLICANT:  Heinonen, Paula
;  APPLICANT:  Alhopuro, Pia
;  APPLICANT:  Karvonen, Matti
;  APPLICANT:  Koulu, Markku
;  APPLICANT:  Pesonen, Ullamari
;  APPLICANT:  Scheinin, Mika
;  APPLICANT:  Salonen, Jukka T
;  APPLICANT:  Tuomainen, Tomi-Pekka
;  APPLICANT:  Lakka, Timo A
;  APPLICANT:  Nyyss"nen, Kristiina
;  APPLICANT:  Salonen, Riitta
;  APPLICANT:  Kauhanen, Jussi
;  APPLICANT:  Valkonen, Veli-Pekka
;  TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
;  TITLE OF INVENTION:  protein, and uses thereof
;  FILE REFERENCE: Alpha-2B-AR variant
;  CURRENT APPLICATION NUMBER: US/09/825,923
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/422,985
;  PRIOR FILING DATE: 2000-05-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1350)
;   OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor
;   OTHER INFORMATION: protein
US-09-825-923-3

  Query Match             100.0%;  Score 21;  DB 9;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209
```

```
RESULT 6
US-10-077-870-3
; Sequence 3, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
;  APPLICANT: Salonen, Jukka T
;  TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
;  FILE REFERENCE: 0933-0183P
;  CURRENT APPLICATION NUMBER: US/10/077,870
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: FI 20010323
;  PRIOR FILING DATE: 2001-02-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1350)
;   OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor protein
US-10-077-870-3

  Query Match             100.0%;  Score 21;  DB 15;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209


RESULT 7
US-10-001-073-1
; Sequence 1, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
;  APPLICANT: Liggett, Stephen
;  APPLICANT:  Small, Kirsten
;  TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
;  FILE REFERENCE: 13073-PCT
;  CURRENT APPLICATION NUMBER: US/10/001,073
;  CURRENT FILING DATE:  2001-11-01
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-001-073-1

  Query Match             100.0%;  Score 21;  DB 15;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 3.8;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209


RESULT 8
US-10-305-720-1181
; Sequence 1181, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
;  APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
;  TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressio
;  FILE REFERENCE: PA-0002-1 CON
;  CURRENT APPLICATION NUMBER: US/10/305,720
;  CURRENT FILING DATE:  2002-11-26
;  PRIOR APPLICATION NUMBER: 09/016,434
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 1490
;  SOFTWARE: PERL Program
; SEQ ID NO 1181
;   LENGTH: 2072
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No. US20040010136A1 g178197
US-10-305-720-1181

  Query Match             100.0%;  Score 21;  DB 12;  Length 2072;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        601 GCTCATCATCCCTTTCTCGCT 621


RESULT 9
US-10-225-567A-41
; Sequence 41, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
;   LENGTH: 3274
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-41

  Query Match             100.0%;  Score 21;  DB 15;  Length 3274;
  Best Local Similarity   100.0%;  Pred. No. 4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db        189 GCTCATCATCCCTTTCTCGCT 209


RESULT 10
US-10-311-455-48/c
; Sequence 48, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determi
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
```

```
; SEQ ID NO 48
;   LENGTH: 6904
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: 821
;   OTHER INFORMATION: n is a or g or c or t
US-10-311-455-48

  Query Match             100.0%;  Score 21;  DB 13;  Length 6904;
  Best Local Similarity   100.0%;  Pred. No. 4.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||||||||
Db       1716 GCTCATCATCCCTTTCTCGCT 1696


RESULT 11
US-10-001-073-24
; Sequence 24, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
;  APPLICANT: Liggett, Stephen
;  APPLICANT:  Small, Kirsten
;  TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
;  FILE REFERENCE: 13073-PCT
;  CURRENT APPLICATION NUMBER: US/10/001,073
;  CURRENT FILING DATE:  2001-11-01
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 1350
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-001-073-24

  Query Match             92.4%;  Score 19.4;  DB 15;  Length 1350;
  Best Local Similarity   95.2%;  Pred. No. 21;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              ||||  ||||||||||||||||
Db        252 GCTCGTCATCCCTTTCTCGCT 272


RESULT 12
US-10-001-073-25
; Sequence 25, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
;  APPLICANT: Liggett, Stephen
;  APPLICANT:  Small, Kirsten
;  TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
;  FILE REFERENCE: 13073-PCT
;  CURRENT APPLICATION NUMBER: US/10/001,073
;  CURRENT FILING DATE:  2001-11-01
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
;   LENGTH: 1350
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-001-073-25

  Query Match             92.4%;  Score 19.4;  DB 15;  Length 1350;
  Best Local Similarity   95.2%;  Pred. No. 21;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              ||||  ||||||||||||||||
Db        252 GCTCGTCATCCCTTTCTCGCT 272


RESULT 13
US-10-305-720-1180
; Sequence 1180, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
;  APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
;  TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressio
;  FILE REFERENCE: PA-0002-1 CON
;  CURRENT APPLICATION NUMBER: US/10/305,720
;  CURRENT FILING DATE:  2002-11-26
;  PRIOR APPLICATION NUMBER: 09/016,434
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 1490
;  SOFTWARE: PERL Program
; SEQ ID NO 1180
;   LENGTH: 3604
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No. US20040010136A1 g178195
US-10-305-720-1180

  Query Match             92.4%;  Score 19.4;  DB 12;  Length 3604;
  Best Local Similarity   95.2%;  Pred. No. 22;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              ||||  ||||||||||||||||
Db       2329 GCTCGTCATCCCTTTCTCGCT 2349


RESULT 14
US-10-225-567A-39
; Sequence 39, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
;   LENGTH: 3653
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-39

  Query Match             92.4%;  Score 19.4;  DB 15;  Length 3653;
  Best Local Similarity   95.2%;  Pred. No. 22;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              ||||  ||||||||||||||||
Db       1131 GCTCGTCATCCCTTTCTCGCT 1151


RESULT 15
US-10-311-455-46/c
; Sequence 46, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: OLEK, Alexander
;  APPLICANT:  PIEPENBROCK, Christian
;  APPLICANT:  BERLIN, Kurt
;  TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
;  TITLE OF INVENTION:  cytosine methylation
;  FILE REFERENCE: 5013.1014
;  CURRENT APPLICATION NUMBER: US/10/311,455
;  CURRENT FILING DATE:  2002-12-16
;  PRIOR APPLICATION NUMBER: PCT/EP01/07537
;  PRIOR FILING DATE: 2001-07-02
;  PRIOR APPLICATION NUMBER: DE 10032529.7
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: DE 10043826.1
;  PRIOR FILING DATE: 2000-09-01
;  NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 46
;   LENGTH: 7353
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-46

  Query Match             92.4%;  Score 19.4;  DB 13;  Length 7353;
  Best Local Similarity   95.2%;  Pred. No. 22;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| ||||||||||||||||
Db       2102 GCTCGTCATCCCTTTCTCGCT 2082


Search completed: February 12, 2004, 06:15:35
Job time : 124.341 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 03:54:41 ; Search time 1135.79 Seconds
(without alignments)
449.373 Million cell updates/sec

Title: US-09-692-077D-13
Perfect score: 21
Sequence: 1 gctcatcatccctttctcgct 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	19.4	92.4	297	13	BQ302172	BQ302172 QV0-BT026
2	19.4	92.4	2410	11	BC035047	BC035047 Homo sapi
3	18.4	87.6	619	10	AW916723	AW916723 EST348027
c 4	18.4	87.6	658	28	BH735033	BH735033 BOMDM37TR
5	18.4	87.6	676	28	AQ448022	AQ448022 mgxb0016A
c 6	18.4	87.6	839	28	BH573939	BH573939 BOGVT08TR
7	18.4	87.6	909	29	BZ752626	BZ752626 PUDCT72TB
c 8	18.4	87.6	945	29	BZ752631	BZ752631 PUDCT72TD
c 9	17.8	84.8	458	9	AJ514080	AJ514080 AJ514080
c 10	17.8	84.8	505	29	BZ374377	BZ374377 ie24c01.g
c 11	17.8	84.8	642	13	BU100970	BU100970 WHE3360_A
c 12	17.8	84.8	650	10	BE427478	BE427478 PSR7068 I
c 13	17.8	84.8	718	13	BU101040	BU101040 WHE3361_A
14	17.8	84.8	817	9	AI113972	AI113972 GH10539.5
15	17.8	84.8	971	10	BE420874	BE420874 HWM003.D1
c 16	17.8	84.8	2225	12	BG842980	BG842980 1024001A1
c 17	17.4	82.9	260	9	AI978394	AI978394 RZ394.F E
18	17.4	82.9	260	9	AI978395	AI978395 RZ394.R E
c 19	17.4	82.9	283	9	AU165754	AU165754 AU165754
20	17.4	82.9	287	10	BF763943	BF763943 CM0-CS004
c 21	17.4	82.9	416	12	BI798818	BI798818 H117H05 E
22	17.4	82.9	509	28	AQ879807	AQ879807 HS_4821_A
c 23	17.4	82.9	564	12	BI798125	BI798125 H098C04 E
24	17.4	82.9	606	13	BU212727	BU212727 604158457
25	17.4	82.9	665	14	CB664045	CB664045 OSJNEd09I
c 26	17.4	82.9	689	28	BH984874	BH984874 oej72c04.
27	17.4	82.9	706	13	BU310902	BU310902 603544065
28	17.4	82.9	736	28	AZ331421	AZ331421 1M0059C22
29	17.4	82.9	791	14	CB650823	CB650823 OSJNEb15H
30	17.4	82.9	811	29	BZ627353	BZ627353 ih51h11.g
31	17.4	82.9	1369	29	BZ580030	BZ580030 msh2_902.
32	16.8	80.0	163	9	AJ503138	AJ503138 AJ503138
33	16.8	80.0	330	10	BB123120	BB123120 BB123120
34	16.8	80.0	439	9	AJ502560	AJ502560 AJ502560
35	16.8	80.0	439	9	AJ503843	AJ503843 AJ503843
36	16.8	80.0	442	29	BZ319189	BZ319189 hx54d09.g
37	16.8	80.0	461	28	AQ864185	AQ864185 nbeb0022L
c 38	16.8	80.0	462	28	AQ147051	AQ147051 HS_2269_A
c 39	16.8	80.0	464	28	AQ024975	AQ024975 EP(2)0844
40	16.8	80.0	467	9	AI998046	AI998046 701672796
c 41	16.8	80.0	518	10	BF016210	BF016210 uy32d09.y
c 42	16.8	80.0	536	13	BQ827804	BQ827804 LL6in2287
43	16.8	80.0	559	28	AZ156636	AZ156636 SP_0045_A
44	16.8	80.0	570	14	CD552619	CD552619 B0346C06-
c 45	16.8	80.0	593	9	AI728961	AI728961 BNLGHi121

ALIGNMENTS

RESULT 1
BQ302172
LOCUS BQ302172 297 bp mRNA linear EST 16-MAY-2002
DEFINITION QV0-BT0263-101299-072-h10 BT0263 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ302172
VERSION BQ302172.1 GI:20817694
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 297)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

```
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=QV0&t2=QV0-BT0263-
            101299-072-h10&t3=1999-12-10&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 26
            High quality sequence stop: 205.
FEATURES                Location/Qualifiers
     source             1. .297
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /dev_stage="Adult"
                        /clone_lib="BT0263"
                        /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                        SmaI; A mini-library was made by cloning products derived
                        from  ORESTES PCR (U.S. Letters Patent application No. 196
                        ,716 -  Ludwig Institute for Cancer Research) profiles
                        into  the pUC 18 vector. Reverse transcription of tissue
                        mRNA and cDNA amplification were performed under low
                        stringency conditions."
BASE COUNT       40 a    109 c     91 g     57 t
ORIGIN

  Query Match             92.4%;  Score 19.4;  DB 13;  Length 297;
  Best Local Similarity   95.2%;  Pred. No. 3.7e+02;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| ||||||||||||||||
Db        262 GCTCGTCATCCCTTTCTCGCT 282


RESULT 2
BC035047
LOCUS       BC035047                2410 bp    mRNA    linear   HTC 23-SEP-2002
DEFINITION  Homo sapiens, Similar to adrenergic, alpha-2A-, receptor, clone
            IMAGE:5266354, mRNA.
ACCESSION   BC035047
VERSION     BC035047.1  GI:23272892
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2410)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) &  Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 73 Row: o Column: 16
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 15718669
            This clone has the following problem: frame shifted.
FEATURES                Location/Qualifiers
     source             1. .2410
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:5266354"
                        /tissue_type="Testis"
                        /clone_lib="NIH_MGC_97"
                        /lab_host="DH10B"
                        /note="Vector: pBluescript"
BASE COUNT       406 a    813 c    721 g    470 t
ORIGIN

  Query Match             92.4%;  Score 19.4;  DB 11;  Length 2410;
  Best Local Similarity   95.2%;  Pred. No. 8.9e+02;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| ||||||||||||||||
Db        465 GCTCGTCATCCCTTTCTCGCT 485


RESULT 3
AW916723
LOCUS       AW916723                619 bp    mRNA    linear   EST 25-MAY-2000
DEFINITION  EST348027 Rat gene index, normalized rat, norvegicus, Bento Soares
            Rattus norvegicus cDNA clone RGIDT95 5' end, mRNA sequence.
ACCESSION   AW916723
VERSION     AW916723.1  GI:8082456
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
            Kerlavage,A.R. and Adams,M.D.
  TITLE     Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat
            Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Lee, NH
            The Institute for Genomic Research
            9712, Medical Center Drive, Rockville, MD 20850, USA
            Tel: (301)-838-3529
            Fax: (301)-838-0208
            Email: nhlee@tigr.org
            This clone is available through the ATCC, contact the ATCC
            tel#703-365-2700 for further information
            Seq primer: M13 Reverse.
FEATURES                Location/Qualifiers
     source             1. .619
                        /organism="Rattus norvegicus"
                        /mol_type="mRNA"
                        /db_xref="taxon:10116"
                        /clone="RGIDT95"
                        /tissue_type="mix - brain, ovary, placenta, kidney, lung,
                        liver, embryo, heart, muscle, spleen"
                        /lab_host="SOLR"
                        /clone_lib="Rat gene index, normalized rat, norvegicus,
                        Bento Soares"
                        /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                        Xho1; Estimated insert size approx.1 kb"
BASE COUNT       105 a    203 c    177 g    134 t
ORIGIN

  Query Match             87.6%;  Score 18.4;  DB 10;  Length 619;
  Best Local Similarity   95.0%;  Pred. No. 1.3e+03;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CTCATCATCCCTTTCTCGCT 21
```

```
                  |||||||||||||||| ||
Db        588 CTCATCATCCCTTTCTCTCT 607


RESULT 4
BH735033/c
LOCUS       BH735033                 658 bp    DNA     linear   GSS 20-FEB-2002
DEFINITION  BOMDM37TR BO_2_3_KB Brassica oleracea genomic clone BOMDM37,
            genomic survey sequence.
ACCESSION   BH735033
VERSION     BH735033.1  GI:18840428
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 658)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished
COMMENT     Other_GSSs: BOMDM37TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .658
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOMDM37"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      211 a     84 c    159 g    204 t
ORIGIN

  Query Match             87.6%;  Score 18.4;  DB 28;  Length 658;
  Best Local Similarity   95.0%;  Pred. No. 1.4e+03;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CTCATCATCCCTTTCTCGCT 21
              |||||||||||||||| |||
Db        396 CTCATCATCCCTTTCTAGCT 377


RESULT 5
AQ448022
LOCUS       AQ448022                 676 bp    DNA     linear   GSS 08-APR-1999
DEFINITION  mgxb0016A14f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
            clone mgxb0016A14f, genomic survey sequence.
ACCESSION   AQ448022
VERSION     AQ448022.1  GI:4577159
KEYWORDS    GSS.
SOURCE      Magnaporthe grisea (anamorph: Pyricularia grisea)
  ORGANISM  Magnaporthe grisea
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
REFERENCE   1  (bases 1 to 676)
  AUTHORS   Yu,Y., Zhu,H., Boyd,C.A., Gaudette,B., Gayle,A., Kingsbury,R.,
            Phillips,K., Sasinowski,M, Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
            Genome
  JOURNAL   Unpublished
COMMENT     Contact: Dean RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson Universiy, Clemson, SC 29634
            Tel: 864 656 5737
            Fax: 864 656 4293
            Email: rdean@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence start: 37
            High quality sequence stop: 300.
FEATURES             Location/Qualifiers
     source          1. .676
                     /organism="Magnaporthe grisea"
                     /mol_type="genomic DNA"
                     /strain="70-15"
                     /db_xref="taxon:148305"
                     /clone="mgxb0016A14f"
                     /tissue_type="Protoplasts"
                     /lab_host="E. coli DH10B"
                     /clone_lib="CUGI Rice Blast BAC Library"
                     /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
                     Rice blast is one of the most devestating fungal diseases
                     of rice  world wide.  It is a filamentous ascomycete with
                     a haploid genome (n=7)  of approximately 40 Mbp.  Rice
                     blast is an important model fungal pathogen  for studying
                     numerous aspects of the fungal-host interaction.   In
                     order to facilitate genome wide analysis, a BAC library
                     containing  9216 clones with an average insert size of 130
                     kbp was constructed.   This library represents greater
                     than 25X genome coverage.  High density  colony filters
                     are available upon request."
BASE COUNT      115 a    204 c    150 g    206 t      1 others
ORIGIN

  Query Match             87.6%;  Score 18.4;  DB 28;  Length 676;
  Best Local Similarity   95.0%;  Pred. No. 1.4e+03;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CTCATCATCCCTTTCTCGCT 21
              |||||||||||||||| |||
Db        632 CTCATCATCCCTTTCTTGCT 651


RESULT 6
BH573939/c
LOCUS       BH573939                 839 bp    DNA     linear   GSS 14-DEC-2001
DEFINITION  BOGVT08TR BOGV Brassica oleracea genomic clone BOGVT08, genomic
            survey sequence.
ACCESSION   BH573939
VERSION     BH573939.1  GI:17825778
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 839)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished
COMMENT     Other_GSSs: BOGVT08TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .839
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
```

/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGVT08"
/clone_lib="BOGV"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 274 a 106 c 200 g 259 t
ORIGIN

Query Match 87.6%; Score 18.4; DB 28; Length 839;
Best Local Similarity 95.0%; Pred. No. 1.5e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CTCATCATCCCTTTCTCGCT 21
||||||||||||||||| |||
Db 271 CTCATCATCCCTTTCTAGCT 252

RESULT 7
BZ752626
LOCUS BZ752626 909 bp DNA linear GSS 10-MAR-2003
DEFINITION PUDCT72TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa157L24, genomic survey sequence.
ACCESSION BZ752626
VERSION BZ752626.1 GI:28904975
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 909)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished
COMMENT Other_GSSs: PUDCT72TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .909
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa157L24"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high CoT selected genomic DNA library"
BASE COUNT 280 a 178 c 169 g 282 t
ORIGIN

Query Match 87.6%; Score 18.4; DB 29; Length 909;
Best Local Similarity 95.0%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CTCATCATCCCTTTCTCGCT 21
||||||||||||||||| ||||
Db 533 CTCATCATCCCTTTCACGCT 552

RESULT 8
BZ752631/c
LOCUS BZ752631 945 bp DNA linear GSS 10-MAR-2003
DEFINITION PUDCT72TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa157L24, genomic survey sequence.
ACCESSION BZ752631
VERSION BZ752631.1 GI:28904980
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 945)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished
COMMENT Other_GSSs: PUDCT72TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1. .945
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa157L24"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high CoT selected genomic DNA library"
BASE COUNT 295 a 183 c 174 g 293 t
ORIGIN

Query Match 87.6%; Score 18.4; DB 29; Length 945;
Best Local Similarity 95.0%; Pred. No. 1.6e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CTCATCATCCCTTTCTCGCT 21
||||||||||||||||| ||||
Db 565 CTCATCATCCCTTTCACGCT 546

RESULT 9
AJ514080/c
LOCUS AJ514080 458 bp mRNA linear EST 14-NOV-2002
DEFINITION AJ514080 Gossypium hirsutum ovule first day the flower opens Gossypium hirsutum cDNA clone suo86d07r3, mRNA sequence.
ACCESSION AJ514080
VERSION AJ514080.1 GI:24977563
KEYWORDS EST.
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE 1 (bases 1 to 458)
AUTHORS Suo,J., Liang,X. and Xue,Y.
TITLE The construction of a cDNA library from the ovule of Gossypium hirsutum and its EST analysis
JOURNAL Unpublished
COMMENT Contact: Suo J
Centre of Developmental Biology
The Institute of Genetics and Developmental Biology CAS
Nanyitiao 3, Zhongguancun, Haidian, Beijing, 100080, China.
FEATURES Location/Qualifiers
source 1. .458
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="Zhongmian12"
/db_xref="taxon:3635"
/clone="suo86d07r3"
/tissue_type="ovule"

```
                     /dev_stage="first day the flower opens"
                     /clone_lib="Gossypium hirsutum ovule first day the flower
                     opens"
BASE COUNT      180 a    139 c    136 g      3 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 458;
  Best Local Similarity   90.5%;  Pred. No. 2.1e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              ||||||| |||||||||| ||
Db         26 GCTCATCTTCCCTTTCTCCCT 6


RESULT 10
BZ374377/c
LOCUS       BZ374377                505 bp    DNA     linear   GSS 26-NOV-2002
DEFINITION  ie24c01.g2 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone
            ie24c01 5', genomic survey sequence.
ACCESSION   BZ374377
VERSION     BZ374377.1  GI:25460894
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 505)
  AUTHORS   Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
            Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
            Zutavern,T., McCombie,W.R. and Martienssen,R.A.
  TITLE     Genomic shotgun sequences from Zea mays (methyl-filtered)
  JOURNAL   Unpublished
COMMENT     Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mccombie@cshl.org
            Plate: ie24  row: c  column: 01
            Seq primer: -21M13UnivRev
            Class: shotgun
            High quality sequence stop: 505.
FEATURES             Location/Qualifiers
     source          1. .505
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ie24c01"
                     /lab_host="DH5a"
                     /clone_lib="WGS-ZmaysF (DH5a methyl filtered)"
                     /note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
                     The vector was digested with XbaI and one nucleotide was
                     added by fill in in the recessive 3' end. The genomic DNA
                     was nebulized, end repaired, adaptor ligated and size
                     fractionated using sephadex. The resulting fragments were
                     between 0.8 and 3 kb and were cloned into the vector (.x/y
                     reads in M13mp19, .b/g reads in pUC19). The same ligation
                     was transformed into DH5a."
BASE COUNT      112 a    146 c    141 g    106 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 29;  Length 505;
  Best Local Similarity   90.5%;  Pred. No. 2.2e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||| |||| ||
Db        478 GCTCATCATCCCTCTCTCTCT 458


RESULT 11
BU100970/c
LOCUS       BU100970                642 bp    mRNA    linear   EST 29-AUG-2002
DEFINITION  WHE3360_A09_A18ZS Chinese Spring aluminum-stressed root tip cDNA
            library Triticum aestivum cDNA clone WHE3360_A09_A18, mRNA
            sequence.
ACCESSION   BU100970
VERSION     BU100970.1  GI:22548769
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Anderson,O.D., Chao,S., Chin,A., Close,T.J., Crossman,C., Gustafson
            ,J.P., Lazo,G.R., Pham,J., Rausch,C.J., Ross,K., Wilson,C. and Woo
            ,J.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Chinese Spring aluminum-stressed root tip cDNA library
  JOURNAL   Unpublished
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE3360_A09_A18"
                     /tissue_type="Root tip at 1.0 to 1.5 mm stage"
                     /dev_stage="Seedling"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Chinese Spring aluminum-stressed root tip cDNA
                     library"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown for four
                     days under hydroponic conditions with 1ppm aluminum, root
                     tips were excised and snap frozen, and total RNA was
                     prepared at University of Missouri in the JP Gustafson lab
                     (Ross). Poly(A) RNA was purified, a cDNA library was made,
                     and the cDNA clones were in vivo excised to give
                     pBluescript SK(-) phagemids in the TJ Close lab (Chin) at
                     the University of California, Riverside. Plasmid DNA
                     preparations and DNA sequencing were performed in the OD
                     Anderson lab (all other authors)."
BASE COUNT      201 a    124 c    176 g    141 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 13;  Length 642;
  Best Local Similarity   90.5%;  Pred. No. 2.4e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||||||||  ||||
Db        397 GCTCATCATCCCTTTTCCGCT 377


RESULT 12
BE427478/c
LOCUS       BE427478                650 bp    mRNA    linear   EST 24-JUL-2000
DEFINITION  PSR7068 ITEC PSR Wheat Pericarp/Testa Library Triticum aestivum
            cDNA clone PSR7068, mRNA sequence.
ACCESSION   BE427478
```

```
VERSION     BE427478.1  GI:9425321
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 650)
  AUTHORS   Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
            ,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
            Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
            Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
            Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
            Sorrells,M., Warburton,M. and Wenzel,G.
  TITLE     International Triticeae EST Cooperative (ITEC): Production of
            Expressed Sequence Tags for Species of the Triticeae
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Bailey P
            Cereals Group, John Innes Centre
            Norfolk, Norwich NR4 7UH UNITED KINGDOM
            Tel: 44 1603 452571 ext. 2587
            Fax: 44 1603 502241
            Email: paul.bailey@bbsrc.ac.uk
            International Triticeae EST Cooperative (ITEC)
            http://wheat.pw.usda.gov/genome.
FEATURES             Location/Qualifiers
     source          1. .650
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Novosibirskaya 67"
                     /db_xref="taxon:4565"
                     /clone="PSR7068"
                     /tissue_type="pericarp/testa"
                     /dev_stage="3:1 mix 10:20 days post anthesis"
                     /clone_lib="ITEC PSR Wheat Pericarp/Testa Library"
                     /note="Vector: Lambda ZAP II (Stratagene)"
BASE COUNT      190 a    122 c    160 g    162 t     16 others
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 650;
  Best Local Similarity   90.5%;  Pred. No. 2.4e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              ||||||||||||||||  ||||
Db        280 GCTCATCATCCCTTTTCCGCT 260


RESULT 13
BU101040/c
LOCUS       BU101040                718 bp    mRNA    linear   EST 29-AUG-2002
DEFINITION  WHE3361_A09_B17ZS Chinese Spring aluminum-stressed root tip cDNA
            library Triticum aestivum cDNA clone WHE3361_A09_B17, mRNA
            sequence.
ACCESSION   BU101040
VERSION     BU101040.1  GI:22548839
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 718)
  AUTHORS   Anderson,O.D., Chao,S., Chin,A., Close,T.J., Crossman,C., Gustafson
            ,J.P., Lazo,G.R., Pham,J., Rausch,C.J., Ross,K., Wilson,C. and Woo
            ,J.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Chinese Spring aluminum-stressed root tip cDNA library
  JOURNAL   Unpublished
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: SK primer.
FEATURES             Location/Qualifiers
     source          1. .718
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE3361_A09_B17"
                     /tissue_type="Root tip at 1.0 to 1.5 mm stage"
                     /dev_stage="Seedling"
                     /lab_host="E. coli SOLR"
                     /clone_lib="Chinese Spring aluminum-stressed root tip cDNA
                     library"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                     Site_1: EcoRI; Site_2: XhoI; Plants were grown for four
                     days under hydroponic conditions with 1ppm aluminum, root
                     tips were excised and snap frozen, and total RNA was
                     prepared at University of Missouri in the JP Gustafson lab
                     (Ross). Poly(A) RNA was purified, a cDNA library was made,
                     and the cDNA clones were in vivo excised to give
                     pBluescript SK(-) phagemids in the TJ Close lab (Chin) at
                     the University of California, Riverside. Plasmid DNA
                     preparations and DNA sequencing were performed in the OD
                     Anderson lab (all other authors)."
BASE COUNT      222 a    138 c    196 g    162 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 13;  Length 718;
  Best Local Similarity   90.5%;  Pred. No. 2.5e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              ||||||||||||||||  ||||
Db        397 GCTCATCATCCCTTTTCCGCT 377


RESULT 14
AI113972
LOCUS       AI113972                817 bp    mRNA    linear   EST 23-APR-2001
DEFINITION  GH10539.5prime GH Drosophila melanogaster head pOT2 Drosophila
            melanogaster cDNA clone GH10539 5 similar to CG12781: FBan0012781
            located on: 2R 59B4-59B4;: 04/10/2001, mRNA sequence.
ACCESSION   AI113972
VERSION     AI113972.2  GI:13757849
KEYWORDS    EST.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 817)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished
COMMENT     On Sep 2, 1998 this sequence version replaced gi:3514775.
            Other_ESTs: GH10539.3prime
            Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003459: arm:2R [17805870,18105620]
            estimated-cyto:59B2-59D2: 04/10/2001
            Plate: GH.105  row: D  column: 3
            High quality sequence stop: 579
            POLYA=No.
```

```
FEATURES             Location/Qualifiers
     source          1. .817
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="GH10539"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="DH5 - alpha"
                     /clone_lib="GH Drosophila melanogaster head pOT2"
                     /note="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
                     XhoI; Sized fractionated cDNAs were directly ligated into
                     pOT2.  Plasmid cDNA library."
BASE COUNT      156 a    213 c    219 g    229 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 817;
  Best Local Similarity   90.5%;  Pred. No. 2.7e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||| || |||||||||||||
Db        407 GCTCCTCCTCCCTTTCTCGCT 427


RESULT 15
BE420874
LOCUS       BE420874                971 bp    mRNA    linear   EST 24-JUL-2000
DEFINITION  HWM003.D11 ITEC HWM Barley Leaf Library Hordeum vulgare subsp.
            vulgare cDNA clone HWM003.D11, mRNA sequence.
ACCESSION   BE420874
VERSION     BE420874.1  GI:9418717
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 971)
  AUTHORS   Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
            ,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
            Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
            Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
            Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
            Sorrells,M., Warburton,M. and Wenzel,G.
  TITLE     International Triticeae EST Cooperative (ITEC): Production of
            Expressed Sequence Tags for Species of the Triticeae
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Herrmann RG
            Botanisches Institut der LMU
            Menzinger Str. 67, D-80638 Munchen GERMANY
            Fax: 49 30 171683
            Email: herrmann@botanik.biologie.uni-muenchen.de
            International Triticeae EST Cooperative (ITEC)
            http://wheat.pw.usda.gov/genome.
FEATURES             Location/Qualifiers
     source          1. .971
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="barke"
                     /db_xref="taxon:112509"
                     /clone="HWM003.D11"
                     /tissue_type="leaf"
                     /dev_stage="14 day old"
                     /clone_lib="ITEC HWM Barley Leaf Library"
                     /note="Vector: pBluescriptSK(-); 850 bp average insert
                     size."
BASE COUNT      206 a    312 c    145 g    308 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 10;  Length 971;
  Best Local Similarity   90.5%;  Pred. No. 2.9e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GCTCATCATCCCTTTCTCGCT 21
              |||||||||| || |||||||
Db        825 GCTCATCATCTCTCTCTCGCT 845


Search completed: February 12, 2004, 06:05:45
Job time : 1147.79 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 01:54:29 ; Search time 615.146 Seconds
(without alignments)
1396.582 Million cell updates/sec

Title: US-09-692-077D-14
Perfect score: 21
Sequence: 1 aaagccccaccatggtcgggt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		21	100.0	21	6	AX350502	AX350502 Sequence
2	c	21	100.0	828	10	S67316	S67316 alpha 2-adr
3	c	21	100.0	885	9	HUMA2C2	M38742 Human alpha
4		21	100.0	1030	9	HSA325747	AJ325747 Homo sapi
5	c	21	100.0	1344	6	AX350490	AX350490 Sequence
6	c	21	100.0	1344	9	AF316895	AF316895 Homo sapi
7	c	21	100.0	1353	6	AX350489	AX350489 Sequence
8	c	21	100.0	2072	6	AR270618	AR270618 Sequence
9	c	21	100.0	2072	9	HUMADRA2RA	M34041 Human alpha
10	c	21	100.0	3274	6	AX548756	AX548756 Sequence
11	c	21	100.0	9842	9	AF005900	AF005900 Homo sapi
12		21	100.0	22842	9	AC092603	AC092603 Homo sapi
13		17.8	84.8	81920	2	AC019708	AC019708 Drosophil
14		17.8	84.8	153648	3	AC008210	AC008210 Drosophil
15	c	17.8	84.8	181905	3	AC008211	AC008211 Drosophil
16		17.8	84.8	213688	3	AE003755	AE003755 Drosophil
17	c	17.8	84.8	259219	2	AC106994	AC106994 Rattus no
18		17.8	84.8	283336	2	AC097749	AC097749 Rattus no
19		17.4	82.9	18496	1	AE001925	AE001925 Deinococc
20	c	17.4	82.9	201844	9	CNS05TCL	AL355102 Human chr
21	c	17	81.0	8607	1	AE013182	AE013182 Thermoana
22		17	81.0	178967	2	AC128772	AC128772 Rattus no
23	c	17	81.0	275605	2	AC123100	AC123100 Rattus no
24	c	16.8	80.0	636	8	CNS01C0S	AL114932 Botrytis
25	c	16.8	80.0	660	8	CNS01AX1	AL113501 Botrytis
26		16.8	80.0	4313	9	AB020667	AB020667 Homo sapi
27		16.8	80.0	68040	2	AC101898	AC101898 Mus muscu
28		16.8	80.0	88883	9	AC107911	AC107911 Homo sapi
29	c	16.8	80.0	91966	2	AC138708	AC138708 Homo sapi
30	c	16.8	80.0	104435	9	AC016292	AC016292 Homo sapi
31		16.8	80.0	107529	2	AC069085	AC069085 Homo sapi
32	c	16.8	80.0	107529	2	AC069085	AC069085 Homo sapi
33	c	16.8	80.0	110865	7	AY129338	AY129338 Mycobacte
34	c	16.8	80.0	114800	9	AC123567	AC123567 Homo sapi
35	c	16.8	80.0	125856	9	HS1187M17	AL121891 Human DNA
36		16.8	80.0	154961	2	AP005884	AP005884 Oryza sat
37		16.8	80.0	157722	10	AL953890	AL953890 Mouse DNA
38		16.8	80.0	162808	2	AC021340	AC021340 Homo sapi
39	c	16.8	80.0	172945	9	AC108099	AC108099 Homo sapi
40		16.8	80.0	174099	2	AP005789	AP005789 Oryza sat
41		16.8	80.0	176843	2	AC026655	AC026655 Homo sapi
42		16.8	80.0	193641	10	AC121972	AC121972 Mus muscu
43		16.8	80.0	207461	2	AC111108	AC111108 Mus muscu
44		16.8	80.0	207661	9	AC015884	AC015884 Homo sapi
45		16.8	80.0	238129	2	AC094595	AC094595 Rattus no

ALIGNMENTS

RESULT 1
AX350502
LOCUS AX350502 21 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 14 from Patent WO0179561.
ACCESSION AX350502
VERSION AX350502.1 GI:18616098
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liggett,S.B. and Small,K.M.
TITLE Alpha-2 adrenergic receptor polymorphisms
JOURNAL Patent: WO 0179561-A 14 25-OCT-2001;

```
            Liggett, Stephen B. (US) ; Small, Kersten M. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT        5 a      7 c      6 g      3 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 6;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 24;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db          1 AAAGCCCCACCATGGTCGGGT 21


RESULT 2
S67316/c
LOCUS       S67316                828 bp    mRNA    linear   ROD 21-APR-2003
DEFINITION  alpha 2-adrenergic receptor [rats, Sprague-Dawley, islets of
            Langerhans, mRNA Partial, 828 nt].
ACCESSION   S67316
VERSION     S67316.1  GI:456949
KEYWORDS    .
SOURCE      Rattus sp.
  ORGANISM  Rattus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 828)
  AUTHORS   Wang,S.Y. and Pilkey,D.T.
  TITLE     Identification in islets of Langerhans of a new rat alpha
            2-adrenergic receptor
  JOURNAL   Diabetes 43 (1), 127-136 (1994)
  MEDLINE   94085695
   PUBMED   8262309
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 140730] from the original journal article.
            This sequence comes from Fig. 1.
FEATURES             Location/Qualifiers
     source          1. .828
                     /organism="Rattus sp."
                     /mol_type="mRNA"
                     /db_xref="taxon:10118"
     CDS             <1. .>828
                     /note="This sequence comes from Fig. 1"
                     /codon_start=1
                     /product="alpha 2-adrenergic receptor"
                     /protein_id="AAP14006.1"
                     /db_xref="GI:30027780"
                     /translation="CTSSIVHLCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWLI
                     AAVISLPPLIYKGDQGPQPRGRPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIYL
                     IAKRSNRRGPRAKGGPGQGESKQPRPDHGGALASAKLPALASVASAREVNGHSKSTGE
                     KEEGETPEDTGTRALPASWAALPNSGQGQKEGVCGGSPEDEAEEEEEEEEECEPQAVP
                     VSPASACSPPLQQPQGSRVLATLRGQVLLGRGVGAIGGQWWRRRAQLTREKRFTFVLA
                     "
BASE COUNT      148 a    271 c    269 g    140 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 10;  Length 828;
  Best Local Similarity  100.0%;  Pred. No. 15;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        401 AAAGCCCCACCATGGTCGGGT 381


RESULT 3
HUMA2C2/c
LOCUS       HUMA2C2                885 bp    DNA     linear   PRI 06-MAR-1995
DEFINITION  Human alpha-2 adrenergic receptor (ADRA2C) gene, partial cds.
ACCESSION   M38742
VERSION     M38742.1  GI:177867
KEYWORDS    alpha-2 adrenergic receptor.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   Chang,A.C., Ho,T.F. and Chang,N.C.
  TITLE     In vitro amplification by polymerase chain reaction of a partial
            gene encoding the third subtype of alpha-2 adrenergic receptor in
            humans
  JOURNAL   Biochem. Biophys. Res. Commun. 172 (2), 817-823 (1990)
  MEDLINE   91054503
   PUBMED   2173582
COMMENT     Original source text: Human adult neuroblastoma DNA, clone PCRA2.
            Draft entry and computer-readable sequence for [Unpublished (1990)]
            kindly submitted
            by A.C.Chang, 20-SEP-1990.
               Institute of Neuroscience
               155, Sect II, Li-Noon St.
               Taipei, Taiwan, 11221 ROC.
FEATURES             Location/Qualifiers
     source          1. .885
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="4p16.3-p15"
                     /clone="PCRA2"
                     /cell_line="neuroblastoma SK-N-SH"
                     /cell_type="neuroblastoma"
                     /tissue_type="neural"
                     /dev_stage="adult"
     gene            1. .885
                     /gene="ADRA2C"
     CDS             1. .885
                     /partial
                     /gene="ADRA2C"
                     /codon_start=1
                     /product="alpha-2-adrenergic receptor"
                     /protein_id="AAA62823.1"
                     /db_xref="GI:177868"
                     /db_xref="GDB:G00-120-540"
                     /translation="FCTSSIVHLCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWL
                     IAAVISLPPLIYKGDQGPQPRGRPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIY
                     LIAKRSNRRGPRAKGGPGQGESKQPRPDHGGALASAKLPALASVASAREVNGHSKSTG
                     EKEEGETPEDTGTRALPPSWAALPNSGQGQKEGVCGASPEDEAEEEEEEEEEEECEP
                     QAVPVSPASACSPPLQQPQGSRVLATLRGQVLLGRGVGAIGGQWWRRRAQLTREKRFT
                     FVLAVVIGVFVLCWFPFFF"
BASE COUNT      152 a    285 c    285 g    163 t
ORIGIN

  Query Match           100.0%;  Score 21;  DB 9;  Length 885;
  Best Local Similarity  100.0%;  Pred. No. 15;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        404 AAAGCCCCACCATGGTCGGGT 384


RESULT 4
HSA325747
LOCUS       HSA325747              1030 bp    DNA     linear   PRI 18-JUL-2002
DEFINITION  Homo sapiens genomic sequence surrounding NotI site, clone
            NB1-818S.
ACCESSION   AJ325747
VERSION     AJ325747.1  GI:15870141
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1030)
  AUTHORS   Kutsenko,A.S., Gizatullin,R.Z., Al-Amin,A.N., Wang,F., Kvasha,S.M.,
            Podowski,R.M., Matushkin,Y.G., Gyanchandani,A., Muravenko,O.V.,
            Levitsky,V.G., Kolchanov,N.A., Protopopov,A.I., Kashuba,V.I.,
            Kisselev,L.L., Wasserman,W., Wahlestedt,C. and Zabarovsky,E.R.
  TITLE     NotI flanking sequences: a tool for gene discovery and verification
            of the human genome
  JOURNAL   Nucleic Acids Res. 30 (14), 3163-3170 (2002)
  MEDLINE   22131767
   PUBMED   12136098
REFERENCE   2  (bases 1 to 1030)
  AUTHORS   Zabarovsky,E.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAY-2001) Microbiology and Tumorbiology Centre,
            Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .1030
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="NB1-818S"
BASE COUNT      211 a    329 c    308 g    182 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 9;  Length 1030;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        393 AAAGCCCCACCATGGTCGGGT 413


RESULT 5
AX350490/c
LOCUS       AX350490               1344 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 2 from Patent WO0179561.
ACCESSION   AX350490
VERSION     AX350490.1  GI:18616092
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liggett,S.B. and Small,K.M.
  TITLE     Alpha-2 adrenergic receptor polymorphisms
  JOURNAL   Patent: WO 0179561-A 2 25-OCT-2001;
            Liggett, Stephen B. (US) ; Small, Kersten M. (US)
FEATURES             Location/Qualifiers
     source          1. .1344
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      220 a    458 c    400 g    266 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 6
AF316895/c
LOCUS       AF316895               1344 bp    DNA     linear   PRI 13-FEB-2001
DEFINITION  Homo sapiens alpha 2B adrenergic receptor (ADRA2B) gene, complete
            cds.
ACCESSION   AF316895
VERSION     AF316895.1  GI:12698669
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1344)
  AUTHORS   Small,K.M., Brown,K.M., Forbes,S.L. and Liggett,S.B.
  TITLE     Polymorphic deletion of three intracellular acidic residues of the
            alpha 2B-adrenergic receptor decreases G protein-coupled receptor
            kinase-mediated phosphorylation and desensitization
  JOURNAL   J. Biol. Chem. 276 (7), 4917-4922 (2001)
  MEDLINE   21265012
   PUBMED   11056163
REFERENCE   2  (bases 1 to 1344)
  AUTHORS   Small,K.M., Brown,K.M., Forbes,S.L. and Liggett,S.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2000) Internal Medicine, University of
            Cincinnati, 231 Bethesda Ave, Cincinnati, OH 45267, USA
FEATURES             Location/Qualifiers
     source          1. .1344
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
     gene            <1. .>1344
                     /gene="ADRA2B"
     mRNA            <1. .>1344
                     /gene="ADRA2B"
                     /product="alpha 2B adrenergic receptor"
     CDS             1. .1344
                     /gene="ADRA2B"
                     /note="G-protein coupled receptor"
                     /codon_start=1
                     /product="alpha 2B adrenergic receptor"
                     /protein_id="AAK01635.1"
                     /db_xref="GI:12698670"
                     /translation="MDHQDPYSVQATAAIAAAITFLILFTIFGNALVILAVLTSRSLR
                     APQNLFLVSLAAADILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVH
                     LCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPRG
                     RPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIYLIAKRSNRRGPRAKGGPGQGES
                     KQPRPDHGGALASAKLPALASVASAREVNGHSKSTGEKEEGETPEDTGTRALPPSWAA
                     LPNSGQGQKEGVCGASPEDEAEEEEEEEECEPQAVPVSPASACSPPLQQPQGSRVLA
                     TLRGQVLLGRGVGAIGGQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGA
                     ICPKHCKVPHGLFQFFFWIGYCNSSLNPVIYTIFNQDFRRAFRRILCRPWTQTAW"
     variation       891^892
                     /gene="ADRA2B"
                     /note="compared to wild type sequence presented in GenBank
                     Accession Number AF005900; polymorphic sequence lacks
                     three glu residues at this location"
                     /frequency="Caucasians 0.31; African-Americans 0.12"
                     /replace="gaagaggag"
BASE COUNT      220 a    458 c    400 g    266 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 9;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 7
AX350489/c
LOCUS       AX350489               1353 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 1 from Patent WO0179561.
ACCESSION   AX350489
VERSION     AX350489.1  GI:18616091
KEYWORDS    .
SOURCE      Homo sapiens (human)
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liggett,S.B. and Small,K.M.
  TITLE     Alpha-2 adrenergic receptor polymorphisms
  JOURNAL   Patent: WO 0179561-A 1 25-OCT-2001;
            Liggett, Stephen B. (US) ; Small, Kersten M. (US)
FEATURES             Location/Qualifiers
     source          1. .1353
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      224 a    458 c    405 g    266 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 8
AR270618/c
LOCUS       AR270618                2072 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 1181 from patent US 6500938.
ACCESSION   AR270618
VERSION     AR270618.1  GI:29701852
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2072)
  AUTHORS   Au-Young,J. and Seilhamer,J.J.
  TITLE     Composition for the detection of signaling pathway gene expression
  JOURNAL   Patent: US 6500938-A 1181 31-DEC-2002;
FEATURES             Location/Qualifiers
     source          1. .2072
                     /organism="unknown"
BASE COUNT      316 a    705 c    660 g    391 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 2072;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db       1098 AAAGCCCCACCATGGTCGGGT 1078


RESULT 9
HUMADRA2RA/c
LOCUS       HUMADRA2RA              2072 bp    DNA     linear   PRI 30-OCT-1994
DEFINITION  Human alpha-2-adrenergic receptor (alpha-2 c2) gene, complete cds.
ACCESSION   M34041
VERSION     M34041.1  GI:178197
KEYWORDS    alpha-2-adrenergic receptor; plasma membrane protein;
            receptor-coupled G protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2072)
  AUTHORS   Lomasney,J.W., Lorenz,W., Allen,L.F., King,K., Regan,J.W.,
            Yang-Feng,T.L., Caron,M.G. and Lefkowitz,R.J.
  TITLE     Expansion of the alpha 2-adrenergic receptor family: cloning and
            characterization of a human alpha 2-adrenergic receptor subtype,
            the gene for which is located on chromosome 2
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 87 (13), 5094-5098 (1990)
  MEDLINE   90311349
   PUBMED   2164221
COMMENT     Original source text: Human placenta DNA, clone alpha-2 C2.
            Draft entry and computer-readable sequence for [1] kindly submitted
            by J.W.Lomasney, 03-MAY-1990, for release after publication.
FEATURES             Location/Qualifiers
     source          1. .2072
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="2"
     gene            413. .1765
                     /gene="ADRA2B"
     CDS             413. .1765
                     /gene="ADRA2B"
                     /note="alpha-2-adrenergic receptor (alpha-2 C2)  old gene
                     name 'ADRA2RL1'"
                     /codon_start=1
                     /protein_id="AAA51666.1"
                     /db_xref="GI:178198"
                     /db_xref="GDB:G00-120-539"
                     /translation="MDHQDPYSVQATAAIAAAITFLILFTIFGNALVILAVLTSRSLR
                     APQNLFLVSLAAADILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVH
                     LCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPRG
                     RPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIYLIAKRSNRRGPRAKGGPGQGES
                     KQPRPDHGGALASAKLPALASVASAREVNGHSKSTGEKEEGETPEDTGTRALPPSWAA
                     LPNSGQGQKEGVCGASPEDEAEEEEEEEEEEECEPQAVPVSPASACSPPLQQPQGSR
                     VLATLRGQVLLGRGVGAIGGQWWRRRAHVTREKRFTFVLAVVIGVFVLCWFPFFFSYS
                     LGAICPKHCKVPHGLFQFFFWIGYCNSSLNPVIYTIFNQDFRRAFRRILCRPWTQTAW
                     "
BASE COUNT      316 a    705 c    660 g    391 t
ORIGIN      Chromosome 2.

  Query Match             100.0%;  Score 21;  DB 9;  Length 2072;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db       1098 AAAGCCCCACCATGGTCGGGT 1078


RESULT 10
AX548756/c
LOCUS       AX548756                3274 bp    DNA     linear   PAT 26-NOV-2002
DEFINITION  Sequence 41 from Patent WO02061087.
ACCESSION   AX548756
VERSION     AX548756.1  GI:25813686
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burmer,G.C., Roush,C.L. and Brown,J.P.
  TITLE     Antigenic peptides, such as for G protein-coupled receptors
            (GPCRs), antibodies thereto, and systems for identifying such
            antigenic peptides
  JOURNAL   Patent: WO 02061087-A 41 08-AUG-2002;
            Lifespan Biosciences, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .3274
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      587 a    979 c    967 g    741 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 6;  Length 3274;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
```

```
                    ||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 11
AF005900/c
LOCUS       AF005900                9842 bp    DNA     linear   PRI 08-JUL-1997
DEFINITION  Homo sapiens alpha2B-adrenergic receptor (alpha2C2AR) gene,
            complete cds.
ACCESSION   AF005900
VERSION     AF005900.1  GI:2245627
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9842)
  AUTHORS   Cayla,C., Schaak,S., Bouloumie,A., Devedjian,J.C. and Paris,H.
  TITLE     Alpha2C2-adrenergic receptor gene
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 9842)
  AUTHORS   Cayla,C., Schaak,S., Bouloumie,A., Devedjian,J.C. and Paris,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-1997) INSERM Unit 317, Institut Louis Bugnard,
            CHU Rangueil, Toulouse 31403, France
FEATURES             Location/Qualifiers
     source          1. .9842
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone_lib="brain fetal genomic library"
     gene            5398. .6750
                     /gene="alpha2C2AR"
                     /note="alpha2C2 adrenergic receptor gene"
     CDS             5398. .6750
                     /gene="alpha2C2AR"
                     /note="adrenergic receptor"
                     /codon_start=1
                     /product="alpha2B-adrenergic receptor"
                     /protein_id="AAB62558.1"
                     /db_xref="GI:2245628"
                     /translation="MDHQDPYSVQATAAIAAAITFLILFTIFGNALVILAVLTSRSLR
                     APQNLFLVSLAAADILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVH
                     LCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPRG
                     RPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIYLIAKRSNRRGPRAKGGPGQGES
                     KQPRPDHGGALASAKLPALASVASAREVNGHSKSTGEKEEGETPEDTGTRALPPSWAA
                     LPNSGQGQKEGVCGASPEDEAEEEEEEEEEEECEPQAVPVSPASACSPPLQQPQGSR
                     VLATLRGQVLLGRGVGAIGGQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYS
                     LGAICPKHCKVPHGLFQFFFWIGYCNSSLNPVIYTIFNQDFRRAFRRILCRPWTQTAW
                     "
     polyA_signal    8649. .8654
BASE COUNT     1870 a   2747 c   3225 g   1998 t      2 others
ORIGIN

  Query Match           100.0%;  Score 21;  DB 9;  Length 9842;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db       6083 AAAGCCCCACCATGGTCGGGT 6063


RESULT 12
AC092603
LOCUS       AC092603               22842 bp    DNA     linear   PRI 01-MAR-2002
DEFINITION  Homo sapiens BAC clone RP11-139J6 from 2, complete sequence.
ACCESSION   AC092603 AC073396
VERSION     AC092603.2  GI:16303539
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 22842)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 22842)
  AUTHORS   Martinka,S., Abbott,A., Hawkins,M., Elliott,G. and Doebber,A.
  TITLE     The sequence of Homo sapiens BAC clone RP11-139J6
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 22842)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 22842)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 22842)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 22842)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Oct 20, 2001 this sequence version replaced gi:14916188.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0139J06
                Drafting Center: WIBR
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
```

```
          VECTOR:  pBACe3.6

          NEIGHBORING SEQUENCE INFORMATION:
          The clone sequenced to the left is RP11-401C13, 2000 bp overlap;
          the clone sequenced to the right is RP11-574O17, 2000 bp overlap.
          Actual end of this clone is at base position 48999 of RP11-574O17.

          Polymorphisms have been identified between AC013272 and AC092603.

          The sequence of AC073396 has been incorporated into AC092603.
FEATURES             Location/Qualifiers
     source          1. .22842
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-139J6"
                     /clone_lib="RPCI-11"
     repeat_region   352. .465
                     /rpt_family="MIR"
     repeat_region   1377. .1424
                     /rpt_family="(TG)n"
     repeat_region   1424. .1449
                     /rpt_family="(GA)n"
     repeat_region   1495. .1665
                     /rpt_family="Alu"
     repeat_region   1669. .1966
                     /rpt_family="Alu"
     repeat_region   2812. .2918
                     /rpt_family="MIR"
     repeat_region   3183. .3448
                     /rpt_family="CT-rich"
     repeat_region   5541. .5566
                     /rpt_family="AT_rich"
     repeat_region   5576. .5870
                     /rpt_family="Alu"
     repeat_region   5873. .6183
                     /rpt_family="Alu"
     repeat_region   6195. .6301
                     /rpt_family="MIR"
     repeat_region   6627. .6733
                     /rpt_family="L2"
     repeat_region   7278. .7415
                     /rpt_family="MIR"
     repeat_region   7698. .7723
                     /rpt_family="(GGCTG)n"
     repeat_region   8140. .8218
                     /rpt_family="G-rich"
     repeat_region   8863. .8922
                     /rpt_family="MIR"
     repeat_region   9191. .9346
                     /rpt_family="MIR"
     repeat_region   10172. .10296
                     /rpt_family="MIR"
     repeat_region   10436. .10550
                     /rpt_family="MIR"
     repeat_region   11016. .11169
                     /rpt_family="MIR"
     repeat_region   11617. .11882
                     /rpt_family="L2"
     misc_feature    12062. .12532
                     /note="match to EST AI733093 (NID:g5054206) oj40h05.x5"
     misc_feature    12062. .12257
                     /note="match to EST AA887330 (NID:g3002438) oj40h05.s1"
     misc_feature    12378. .12902
                     /note="match to EST AI791589 (NID:g5339305) oj40h05.y5"
     misc_feature    13059
                     /note="match to EST BF475329 (NID:g11546156)"
     misc_feature    13072. .13459
                     /note="match to EST BF475329 (NID:g11546156)"
     misc_feature    14030. .14244
                     /note="similar to Homo sapiens EST T39448 (NID:g647179)
                     ya06a09.r2"
     misc_feature    14060. .14247
                     /note="similar to Homo sapiens EST AL549866
                     (NID:g12886265)"
     misc_feature    14060. .14247
                     /note="similar to Homo sapiens EST AL573897
                     (NID:g12933576)"
     repeat_region   14413. .14439
                     /rpt_family="(TCC)n"
     misc_feature    14762. .14822
                     /note="similar to Homo sapiens EST AL544609
                     (NID:g12877089)"
     misc_feature    14849. .16413
                     /note="CpG_island (%GC=74.2, o/e=0.90, #CpGs=191)"
     misc_feature    14900. .15218
                     /note="similar to Homo sapiens EST AL544609
                     (NID:g12877089)"
     misc_feature    15084. .15487
                     /note="similar to Sus scrofa EST AW785035 (NID:g7841811)"
     misc_feature    15110. .15648
                     /note="similar to Rattus norvegicus EST AW916723
                     (NID:g8082456)"
     repeat_region   15747. .15832
                     /rpt_family="G-rich"
     repeat_region   18481. .18596
                     /rpt_family="L2"
     repeat_region   19266. .19675
                     /rpt_family="L2"
     repeat_region   20027. .20223
                     /rpt_family="MER1_type"
     repeat_region   20316. .20416
                     /rpt_family="L2"
     repeat_region   20537. .20973
                     /rpt_family="L2"
     misc_feature    22057. .22471
                     /note="match to EST AA830781 (NID:g2903880) oc56b01.s1"
     misc_feature    22057. .22208
                     /note="match to EST AA836522 (NID:g2910841) od22d08.s1"
BASE COUNT     4797 a   7173 c   6248 g   4624 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 9;  Length 22842;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db      14645 AAAGCCCCACCATGGTCGGGT 14665


RESULT 13
AC019708
LOCUS       AC019708               81920 bp    DNA     linear   HTG 03-JAN-2000
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AC019708
VERSION     AC019708.1  GI:6665189
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 81920)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10210696 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
```

```
FEATURES             Location/Qualifiers
     source          1. .81920
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
BASE COUNT    24542 a  16012 c  16326 g  25040 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 2;  Length 81920;
  Best Local Similarity   90.5%;  Pred. No. 3.3e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              ||||||||||| | |||||||
Db      29462 AAAGCCCCACCTTTGTCGGGT 29482


RESULT 14
AC008210
LOCUS       AC008210              153648 bp    DNA     linear   INV 27-FEB-2001
DEFINITION  Drosophila melanogaster, chromosome 3R, region 97A-97A, BAC clone
            BACR08G22, complete sequence.
ACCESSION   AC008210
VERSION     AC008210.4  GI:13129412
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 153648)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome 3R, region 97A-97A
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 153648)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUL-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Feb 27, 2001 this sequence version replaced gi:6633919.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .153648
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="3R"
                     /map="97A-97A"
                     /clone="BACR08G22 (D774)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
BASE COUNT    45062 a  31425 c  31310 g  45851 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 3;  Length 153648;
  Best Local Similarity   90.5%;  Pred. No. 3.1e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              ||||||||||| | |||||||
Db      91650 AAAGCCCCACCTTTGTCGGGT 91670


RESULT 15
AC008211/c
LOCUS       AC008211              181905 bp    DNA     linear   INV 24-FEB-2001
DEFINITION  Drosophila melanogaster, chromosome 3R, region 97A-97A, BAC clone
            BACR08N16, complete sequence.
ACCESSION   AC008211
VERSION     AC008211.9  GI:13122719
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 181905)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome 3R, region 97A-97A
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 181905)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUL-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Feb 24, 2001 this sequence version replaced gi:7208848.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
```

```
           to bdgp@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .181905
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="3R"
                     /map="97A-97A"
                     /clone="BACR08N16 (D775)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
BASE COUNT    55481 a  36378 c  36144 g  53902 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 3;  Length 181905;
  Best Local Similarity   90.5%;  Pred. No. 3e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||| | |||||||
Db     149393 AAAGCCCCACCTTTGTCGGGT 149373


Search completed: February 12, 2004, 04:41:41
Job time : 620.146 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 01:53:19 ; Search time 110.378 Seconds
(without alignments)
513.582 Million cell updates/sec

Title: US-09-692-077D-14
Perfect score: 21
Sequence: 1 aaagccccaccatggtcgggt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21	100.0	21	23	AAI99908	Human alpha-2BAR g
c	2	21	100.0	65	24	ABN32100	Rat spliced transc
c	3	21	100.0	1344	22	AAD04761	Human alpha2B-adre
c	4	21	100.0	1344	23	AAI99906	Human alpha-2BAR t
c	5	21	100.0	1344	24	AAD44388	Human alpha-2B-adr
c	6	21	100.0	1353	22	AAD04762	Human alpha2B-adre
c	7	21	100.0	1353	23	AAI99905	Human alpha-2BAR t
c	8	21	100.0	1353	24	AAD44389	Human alpha-2B-adr
c	9	21	100.0	2064	12	AAQ14151	Human alpha 2 beta
c	10	21	100.0	2064	18	AAT59499	Human alpha-2b adr
c	11	21	100.0	2072	25	ACA56583	Human signalling p
c	12	21	100.0	3274	25	ABZ42624	Human alpha 2b-adr
c	13	17.4	82.9	1755	23	AAS92399	DNA encoding novel
	14	16.8	80.0	447	23	ABV07991	Human prostate exp
	15	16.4	78.1	990	22	ABA12198	Human nervous syst
c	16	16.4	78.1	2262	22	AAI98095	Human neuroblastom
	17	16.4	78.1	6997	25	ABX70470	DNA encoding human
c	18	16.4	78.1	28438	25	AAD47687	Human kinase prote
	19	16.4	78.1	30032	22	ABA17086	Human nervous syst
	20	16.2	77.1	203	25	ABZ19949	Group III cDNA can
c	21	16.2	77.1	1182	19	AAV50488	Streptomyces clavu
	22	16.2	77.1	2216	22	AAK53307	Human polynucleoti
	23	16.2	77.1	2216	22	AAK53308	Human polynucleoti
	24	16.2	77.1	2216	22	AAI61088	Human polynucleoti
c	25	16.2	77.1	2441	22	AAK52324	Human polynucleoti
c	26	16.2	77.1	2582	22	AAK52323	Human polynucleoti
c	27	16.2	77.1	4435	24	ABQ93524	Human cDNA SEQ ID
	28	16.2	77.1	7193	19	AAV50431	Streptomyces clavu
	29	16.2	77.1	34088	23	AAS59566	Propionibacterium
	30	15.8	75.2	51	22	AAL28410	Human SNP oligonuc
	31	15.8	75.2	704	22	AAI85859	Human polynucleoti
	32	15.8	75.2	1139	15	AAQ71655	Sequence encoding
c	33	15.8	75.2	1615	22	AAS28640	Genomic sequence #
c	34	15.8	75.2	1710	22	AAH46866	Human 25278 sulfat
c	35	15.8	75.2	1835	25	ABX72202	Human NOVX polynuc
c	36	15.8	75.2	2634	24	AAD41400	Human NZMS-13 cDNA
	37	15.8	75.2	2754	18	AAT58505	H218 cDNA encoding
	38	15.8	75.2	2754	20	AAV84039	DNA encoding rat H
	39	15.8	75.2	2754	25	ABQ81032	Rat Endothelial Di
c	40	15.8	75.2	2940	22	AAH46862	Human 25278 sulfat
c	41	15.8	75.2	19846	22	AAS36182	Human cardiovascul
c	42	15.8	75.2	19846	22	AAS36183	Human cardiovascul
	43	15.8	75.2	465237	24	ABQ87681	Human oestrogen re
	44	15.8	75.2	465237	24	ABA90193	Human oestrogen re
	45	15.8	75.2	465237	25	ABX33717	Gene encoding huma

ALIGNMENTS

RESULT 1
AAI99908
ID AAI99908 standard; DNA; 21 BP.
XX
AC AAI99908;
XX
DT 18-FEB-2002 (first entry)
XX
DE Human alpha-2BAR genotyping PCR primer SEQ ID NO 14.
XX
KW Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW polymorphic site; allelic variant; cardiovascular disease;
KW central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW phosphorylation; inositol phosphate; alpha-2BAR; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200179561-A2.
XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;

XX
DR WPI; 2001-611728/70.
XX
PT Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT determining whether an individual is at increased risk of developing a
PT disease associated with the corresponding receptor comprises detecting
PT a polymorphic site -
XX
PS Claim 10; Page 112; 163pp; English.
XX
CC The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC alpha2A or alpha2C or fragment or complement of; and
CC (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC positions 961-972 of (III). The method may be used for genotyping an
CC alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC whether an individual is at increased risk of developing a disease
CC associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC polymorphic site which correlate to disease selected from cardiovascular
CC disease, central nervous system disease and combinations of these. In
CC addition, the technique may be used to predict an individual's response
CC to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC these) by detecting the polymorphic site and correlating the site to a
CC predetermined response (where the response is correlated to adenylyl
CC cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC levels). The present sequence is that of a human alpha-2BAR PCR primer,
CC useful for the genotyping methods of the invention.
XX
SQ Sequence 21 BP; 5 A; 7 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 21; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGGT 21
||||||||||||||||||||||
Db 1 AAAGCCCCACCATGGTCGGGT 21

RESULT 2
ABN32100/c
ID ABN32100 standard; DNA; 65 BP.
XX
AC ABN32100;
XX
DT 15-JUL-2002 (first entry)
XX
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:4848.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Rattus norvegicus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX

DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 4848; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 13 A; 22 C; 21 G; 9 T; 0 other;

Query Match 100.0%; Score 21; DB 24; Length 65;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGGT 21
||||||||||||||||||||||
Db 45 AAAGCCCCACCATGGTCGGGT 25

RESULT 3
AAD04761/c
ID AAD04761 standard; DNA; 1344 BP.
XX
AC AAD04761;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human alpha2B-adrenoceptor (alpha2B-AR) variant gene.
XX
KW Human; cardiant; gene therapy; alpha2B-adrenoceptor; alpha2B-AR;
KW glutamic acid repeat; intracellular loop; chromosome 2; catecholamine;
KW norepinephrine; epinephrine; therapy; vascular contraction; variant;
KW coronary artery; coronary heat disease; CHD; chronic angina pectoris;
KW acute myocardial infarction; AMI; Prinzmetal's variant; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1344
FT /*tag= a
FT /product= "Human alpha2B-adrenoceptor (alpha2B-AR)
FT variant protein"
XX
PN WO200129082-A1.
XX
PD 26-APR-2001.
XX

PF 20-OCT-2000; 2000WO-FI00913.
XX
PR 22-OCT-1999; 99US-0422985.
XX
PA (JUVA-) JUVANTIA PHARMA LTD OY.
XX
PI Snapir A, Heinonen P, Alhopuro P, Karvonen M, Koulu M, Pesonen U;
PI Scheinin M, Salonen JT, Tuomainen T, Lakka TA, Nyyssoenen K;
PI Salonen R, Kauhanen J, Valkonen V;
XX
DR WPI; 2001-300318/31.
DR P-PSDB; AAE00989.
XX
PT New DNA molecule encoding variant specific adrenoceptor protein with
PT deletion of specific amino acids located in the third intracellular
PT loop of the polypeptide, for treating vascular contraction of coronary
PT arteries -
XX
PS Claim 3; Page 24-26; 37pp; English.
XX
CC The present sequence is a gene encoding human alpha2B-adrenoceptor
CC (alpha2B-AR) variant protein. Alpha2B-AR has a glutamic acid repeat
CC element (amino acids 298-309) of 12 glutamates, in an acidic stretch of
CC 18 amino acids (amino acids 294-311), located in the third intracellular
CC loop of the receptor polypeptide. The variant is obtained by deletion of
CC three glutamates from the Glu repeat (amino acids 307-309). Alpha2B-AR
CC gene is located on chromosome 2. Alpha2-AR mediate many of the
CC physiological effects of the catecholamines, norepinephrine and
CC epinephrine. An antagonist of alpha2B-adrenoceptor is useful for treating
CC a mammal suffering from vascular contraction of coronary arteries and a
CC disease involving vascular contraction of coronary arteries which is
CC clinically expressed as coronary heat disease (CHD), unstable chronic
CC angina pectoris which is clinically expressed as Prinzmetal's variant
CC form or acute myocardial infarction (AMI). Alpha2B-AR gene is used in
CC gene therapy.
XX
SQ Sequence 1344 BP; 219 A; 459 C; 400 G; 266 T; 0 other;

Query Match 100.0%; Score 21; DB 22; Length 1344;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGGT 21
||||||||||||||||||||
Db 686 AAAGCCCCACCATGGTCGGGT 666

RESULT 4
AAI99906/c
ID AAI99906 standard; DNA; 1344 BP.
XX
AC AAI99906;
XX
DT 18-FEB-2002 (first entry)
XX
DE Human alpha-2BAR third intracellular loop variant encoding DNA.
XX
KW Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW polymorphic site; allelic variant; cardiovascular disease;
KW central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW phosphorylation; inositol phosphate; alpha-2BAR; chromosome 2; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1344
FT /*tag= a
FT /product= "alpha-2BAR"
FT /note= "sequence is deleted for a 9 nucleotide
FT polymorphic site found at nucleotides 901-909
FT of the wildtype alpha-2BAR protein (AAI99905)"
XX
PN WO200179561-A2.
XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;
XX
DR WPI; 2001-611728/70.
DR P-PSDB; AAM52118.
XX
PT Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT determining whether an individual is at increased risk of developing a
PT disease associated with the corresponding receptor comprises detecting
PT a polymorphic site -
XX
PS Claim 5; Page 144-145; 163pp; English.
XX
CC The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC alpha2A or alpha2C or fragment or complement of; and
CC (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC positions 961-972 of (III). The method may be used for genotyping an
CC alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC whether an individual is at increased risk of developing a disease
CC associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC polymorphic site which correlate to disease selected from cardiovascular
CC disease, central nervous system disease and combinations of these. In
CC addition, the technique may be used to predict an individual's response
CC to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC these) by detecting the polymorphic site and correlating the site to a
CC predetermined response (where the response is correlated to adenylyl
CC cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC levels). The present sequence is that of the third intracellular loop of
CC the human alpha-2BAR variant, the sequence is deleted for a 9 nucleotide
CC polymorphic site found at nucleotides 901-909 of the wildtype gene
CC (AAI99905).
XX
SQ Sequence 1344 BP; 220 A; 458 C; 400 G; 266 T; 0 other;

Query Match 100.0%; Score 21; DB 23; Length 1344;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGGT 21
||||||||||||||||||||
Db 686 AAAGCCCCACCATGGTCGGGT 666

RESULT 5
AAD44388/c
ID AAD44388 standard; DNA; 1344 BP.
XX
AC AAD44388;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human alpha-2B-adrenoceptor variant DNA.
XX
KW Human; hypertension; alpha-2B-adrenoceptor; AR; antihypertensive;
KW hypertension; hypotensive; variant; gene; ds.

```
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1344
FT                   /*tag=  a
FT                   /product= "Human alpha-2B-adrenoceptor variant
FT                   protein"
XX
PN   WO200266617-A1.
XX
PD   29-AUG-2002.
XX
PF   13-FEB-2002; 2002WO-FI00113.
XX
PR   20-FEB-2001; 2001FI-0000323.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Salonen J;
XX
DR   WPI; 2002-667063/71.
DR   P-PSDB; AAE26633.
XX
PT   Detecting a risk of hypertension and targeting treatment in a subject
PT   by determining the pattern of alleles encoding a variant
PT   alpha-2-adrenoceptor  -
XX
PS   Disclosure; Page 24-26; 35pp; English.
XX
CC   The invention relates to a method for detecting a risk of hypertension
CC   by determining the pattern of alleles encoding a variant alpha-2B-
CC   adrenoceptor (AR) protein. The methods and compositions of the invention
CC   are useful for detecting risks and targeting treatment for hypertension.
CC   The kit is also useful for selecting for clinical drug trials testing
CC   the antihypertensive effect of compounds. The present sequence is human
CC   alpha-2B-adrenoceptor variant DNA.
XX
SQ   Sequence 1344 BP; 219 A; 459 C; 400 G; 266 T; 0 other;

  Query Match             100.0%;  Score 21;  DB 24;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 6
AAD04762/c
ID   AAD04762 standard; DNA; 1353 BP.
XX
AC   AAD04762;
XX
DT   04-JUL-2001  (first entry)
XX
DE   Human alpha2B-adrenoceptor (alpha2B-AR) gene.
XX
KW   Human; cardiant; gene therapy; alpha2B-adrenoceptor; alpha2B-AR;
KW   glutamic acid repeat; intracellular loop; chromosome 2; catecholamine;
KW   norepinephrine; epinephrine; therapy; vascular contraction;
KW   coronary artery; coronary heat disease; CHD; chronic angina pectoris;
KW   acute myocardial infarction; AMI; Prinzmetal's variant; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1353
FT                   /*tag=  a
FT                   /product= "Human alpha2B-adrenoceptor (alpha2B-AR)
FT                   protein"
XX
PN   WO200129082-A1.
XX
PD   26-APR-2001.
XX
PF   20-OCT-2000; 2000WO-FI00913.
XX
PR   22-OCT-1999;   99US-0422985.
XX
PA   (JUVA-) JUVANTIA PHARMA LTD OY.
XX
PI   Snapir A,  Heinonen P,  Alhopuro P,  Karvonen M,  Koulu M,  Pesonen U;
PI   Scheinin M,  Salonen JT,  Tuomainen T,  Lakka TA,  Nyyssoenen K;
PI   Salonen R,  Kauhanen J,  Valkonen V;
XX
DR   WPI; 2001-300318/31.
DR   P-PSDB; AAE00990.
XX
PT   New DNA molecule encoding variant specific adrenoceptor protein with
PT   deletion of specific amino acids located in the third intracellular
PT   loop of the polypeptide, for treating vascular contraction of coronary
PT   arteries  -
XX
PS   Disclosure; Page 27-29; 37pp; English.
XX
CC   The present sequence is a gene encoding human alpha2B-adrenoceptor
CC   (alpha2B-AR) protein. Alpha2B-AR has a glutamic acid repeat element
CC   (amino acids 298-309) of 12 glutamates, in an acidic stretch of 18 amino
CC   acids (amino acids 294-311), located in the third intracellular loop of
CC   the receptor polypeptide. Alpha2B-AR gene is located on chromosome 2.
CC   Alpha2-AR mediate many of the physiological effects of the
CC   catecholamines, norepinephrine and epinephrine. An antagonist of
CC   alpha2B-adrenoceptor is useful for treating a mammal suffering from
CC   vascular contraction of coronary arteries and a disease involving
CC   vascular contraction of coronary arteries which is clinically expressed
CC   as coronary heat disease (CHD), unstable chronic angina pectoris which is
CC   clinically expressed as Prinzmetal's variant form or acute myocardial
CC   infarction (AMI). Alpha2B-AR gene is used in gene therapy.
XX
SQ   Sequence 1353 BP; 223 A; 459 C; 405 G; 266 T; 0 other;

  Query Match             100.0%;  Score 21;  DB 22;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 7
AAI99905/c
ID   AAI99905 standard; DNA; 1353 BP.
XX
AC   AAI99905;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Human alpha-2BAR third intracellular loop encoding DNA.
XX
KW   Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW   polymorphic site; allelic variant; cardiovascular disease;
KW   central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW   phosphorylation; inositol phosphate; alpha-2BAR;
KW   GenBank Accession AF009500; chromosome 2; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1353
FT                   /*tag=  a
FT                   /product= "alpha-2BAR"
FT                   /note= "sequence includes a 9 nucleotide polymorphic site
```

FT at nucleotides 901-909 absent in the alpha-2BAR
FT variant (AAI99906)"
XX
PN WO200179561-A2.
XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;
XX
DR WPI; 2001-611728/70.
DR P-PSDB; AAM52117.
XX
PT Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT determining whether an individual is at increased risk of developing a
PT disease associated with the corresponding receptor comprises detecting
PT a polymorphic site -
XX
PS Claim 4; Page 144; 163pp; English.
XX
CC The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC alpha2A or alpha2C or fragment or complement of; and
CC (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC positions 961-972 of (III). The method may be used for genotyping an
CC alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC whether an individual is at increased risk of developing a disease
CC associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC polymorphic site which correlate to disease selected from cardiovascular
CC disease, central nervous system disease and combinations of these. In
CC addition, the technique may be used to predict an individual's response
CC to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC these) by detecting the polymorphic site and correlating the site to a
CC predetermined response (where the response is correlated to adenylyl
CC cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC levels). The present sequence is that of the third intracellular loop of
CC the human alpha-2BAR (GenBank Accession AF009500), the sequence includes
CC a 9 nucleotide polymorphic site at nucleotides 901-909, absent in the
CC alpha-2BAR variant (AAI99906).
XX
SQ Sequence 1353 BP; 224 A; 458 C; 405 G; 266 T; 0 other;

Query Match 100.0%; Score 21; DB 23; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGGT 21
||||||||||||||||||||
Db 686 AAAGCCCCACCATGGTCGGGT 666

RESULT 8
AAD44389/c
ID AAD44389 standard; DNA; 1353 BP.
XX
AC AAD44389;
XX
DT 13-DEC-2002 (first entry)
XX
DE Human alpha-2B-adrenoceptor gene.
XX
KW Human; hypertension; alpha-2B-adrenoceptor; AR; antihypertensive;
KW hypertension; hypotensive; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1353
FT /*tag= a
FT /product= "Human alpha-2B-adrenoceptor protein"
XX
PN WO200266617-A1.
XX
PD 29-AUG-2002.
XX
PF 13-FEB-2002; 2002WO-FI00113.
XX
PR 20-FEB-2001; 2001FI-0000323.
XX
PA (JURI-) JURILAB LTD OY.
XX
PI Salonen J;
XX
DR WPI; 2002-667063/71.
DR P-PSDB; AAE26634.
XX
PT Detecting a risk of hypertension and targeting treatment in a subject
PT by determining the pattern of alleles encoding a variant
PT alpha-2-adrenoceptor -
XX
PS Disclosure; Page 27-29; 35pp; English.
XX
CC The invention relates to a method for detecting a risk of hypertension
CC by determining the pattern of alleles encoding a variant alpha-2B-
CC adrenoceptor (AR) protein. The methods and compositions of the invention
CC are useful for detecting risks and targeting treatment for hypertension.
CC The kit is also useful for selecting for clinical drug trials testing
CC the antihypertensive effect of compounds. The present sequence is human
CC alpha-2B-adrenoceptor gene.
XX
SQ Sequence 1353 BP; 223 A; 459 C; 405 G; 266 T; 0 other;

Query Match 100.0%; Score 21; DB 24; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGGT 21
||||||||||||||||||||
Db 686 AAAGCCCCACCATGGTCGGGT 666

RESULT 9
AAQ14151/c
ID AAQ14151 standard; DNA; 2064 BP.
XX
AC AAQ14151;
XX
DT 06-JAN-1992 (first entry)
XX
DE Human alpha 2 beta adrenergic receptor gene.
XX
KW Neurotransmission; adrenaline; epinephrine; NGC-alpha2beta; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 288..1752
FT /*tag= a
XX
PN US5053337-A.
XX
PD 01-OCT-1991.

```
XX
PF   30-OCT-1989;   89US-0428856.
XX
PR   30-OCT-1989;   89US-0428856.
XX
PA   (NEUR-) NEUROGENETIC CORP.
XX
PI   Weinshank RL,  Hartig PR;
XX
DR   WPI; 1991-310087/42.
DR   P-PSDB; AAR14149.
XX
PT   Isolated DNA encoding human adrenergic receptor - for detecting
PT   nucleic acids encoding alpha, 2-beta adrenergic receptor, for
PT   screening drugs.
XX
PS   Claim 1; Fig 2; 15pp; English.
XX
CC   Clone NGC-alpha2beta was isolated from a human spleen genomic
CC   library by screening with a fragment of the human 5-HT1A receptor
CC   gene. The gene can be used to express recombinant receptor protein
CC   which can be used to produce antibodies for inhibition of receptor
CC   function.
XX
SQ   Sequence 2064 BP; 319 A; 696 C; 659 G; 390 T; 0 other;

  Query Match             100.0%;  Score 21;  DB 12;  Length 2064;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db       1084 AAAGCCCCACCATGGTCGGGT 1064


RESULT 10
AAT59499/c
ID   AAT59499 standard; DNA; 2064 BP.
XX
AC   AAT59499;
XX
DT   25-MAR-2003  (updated)
DT   06-MAY-1997  (first entry)
XX
DE   Human alpha-2b adrenergic receptor genomic DNA clone.
XX
KW   Alpha-2b adrenergic receptor; adrenoceptor; adrenaline;
KW   epinephrine; signal transduction; neurotransmitter; ligand; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             288..1751
FT                   /*tag= a
XX
PN   US5595880-A.
XX
PD   21-JAN-1997.
XX
PF   22-OCT-1992;   92US-0965040.
XX
PR   30-OCT-1989;   89US-0428856.
PR   30-MAY-1991;   91US-0707604.
PR   22-OCT-1992;   92US-0965040.
XX
PA   (SYNA-) SYNAPTIC PHARM CORP.
XX
PI   Hartig PR,  Weinshank RL;
XX
DR   WPI; 1997-107576/10.
DR   P-PSDB; AAW11804.
XX
PT   Assay for alpha-2b adrenergic receptor ligands - using membranes of
PT   cells expressing recombinant receptor
XX
PS   Disclosure; Fig 2A-E; 16pp; English.
XX
CC   A genomic DNA clone (AAT59499) codes for human alpha-2B adrenergic
CC   receptor (AAW11804), a member of the rhodopsin-like signal transducer
CC   family.  It was isolated from a human spleen genomic library in the
CC   lambda vector Charon 28 by screening with a 1.6 kb fragment of the
CC   human 5-hydroxytryptamine receptor gene.  Plasmid pNGC-alpha-2B
CC   comprising DNA encoding the alpha-2B adrenoceptor is deposited as
CC   ATCC 68144.  Vectors have been adapted to allow prodn. of alpha-2B
CC   adrenoceptor in bacterial, yeast or mammalian cells; transfected
CC   Ltk- cells, designated L-NGC-alpha-2B, are deposited as ATCC CRL
CC   10275.  Membranes of such cells can used in novel methods to
CC   identify drugs which specifically interact with, and bind to, the
CC   alpha-2B adrenergic receptor.
CC   (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ   Sequence 2064 BP; 319 A; 696 C; 659 G; 390 T; 0 other;

  Query Match             100.0%;  Score 21;  DB 18;  Length 2064;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db       1084 AAAGCCCCACCATGGTCGGGT 1064


RESULT 11
ACA56583/c
ID   ACA56583 standard; cDNA; 2072 BP.
XX
AC   ACA56583;
XX
DT   06-JUN-2003  (first entry)
XX
DE   Human signalling pathway polynucleotide probe SEQ ID NO 1181.
XX
KW   Human; probe; ss; array element; Parkinson's disease;
KW   signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW   immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS   Homo sapiens.
XX
PN   US6500938-B1.
XX
PD   31-DEC-2002.
XX
PF   30-JAN-1998;   98US-0016434.
XX
PR   30-JAN-1998;   98US-0016434.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Au-Young J,  Seilhamer JJ;
XX
DR   WPI; 2003-352189/33.
XX
PT   Combination of polynucleotide probes, useful as array elements in a
PT   microarray for monitoring the expression of a number of target
PT   polynucleotides -
XX
PS   Claim 1; SEQ ID NO 1181; 65pp; English.
XX
CC   The invention relates to a combination which, comprises a number of
CC   polynucleotide probes comprising a sequence selected from one of the 1490
CC   sequences mentioned in the specification. The combination is useful as an
CC   array element in a microarray for monitoring the expression of a number
CC   of target polynucleotides. The microarray is particularly useful in the
CC   diagnosis and treatment of cancer and immunopathology and neuropathology.
CC   The microarray is useful in diagnostics and treatment regimens, drug
CC   discovery and development, toxicological and carcinogenicity studies,
```

CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signaling pathway populations which can be used to diagnose
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC and Parkinson's disease. The present sequence represents a polynucleotide
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.
XX
SQ Sequence 2072 BP; 316 A; 705 C; 660 G; 391 T; 0 other;

Query Match 100.0%; Score 21; DB 25; Length 2072;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGGT 21
||||||||||||||||||||||
Db 1098 AAAGCCCCACCATGGTCGGGT 1078

RESULT 12
ABZ42624/c
ID ABZ42624 standard; DNA; 3274 BP.
XX
AC ABZ42624;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human alpha 2b-adrenoceptor nucleotide SEQ ID NO:41.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US50107.
XX
PR 19-DEC-2000; 2000US-257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
DR P-PSDB; ABP81780.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating
PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
PT cancer or autoimmune diseases -
XX
PS Disclosure; Fig 1; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising:
CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular
CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity
CC or avidity for a particular GPCR. (I) can be used as GPCR modulators and
CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting
CC an antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention.
XX
SQ Sequence 3274 BP; 587 A; 979 C; 967 G; 741 T; 0 other;

Query Match 100.0%; Score 21; DB 25; Length 3274;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGGT 21
||||||||||||||||||||||
Db 686 AAAGCCCCACCATGGTCGGGT 666

RESULT 13
AAS92399/c
ID AAS92399 standard; cDNA; 1755 BP.
XX
AC AAS92399;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #28203.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG28212.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 28203; 103pp; English.

```
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1755 BP; 349 A; 589 C; 545 G; 271 T; 1 other;

  Query Match             82.9%;  Score 17.4;  DB 23;  Length 1755;
  Best Local Similarity   94.7%;  Pred. No. 67;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGG 19
              |||||||| ||||||||||
Db        630 AAAGCCCCCCCATGGTCGG 612


RESULT 14
ABV07991
ID   ABV07991 standard; cDNA; 447 BP.
XX
AC   ABV07991;
XX
DT   13-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 7982.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US05171.
XX
PR   17-FEB-2000; 2000US-183319P.
PR   16-MAR-2000; 2000US-189862P.
PR   25-MAY-2000; 2000US-207454P.
PR   09-JUN-2000; 2000US-211314P.
PR   18-JUL-2000; 2000US-219007P.
PR   13-DEC-2000; 2000US-255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer  -
XX
PS   Claim 1; Page 1278; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for:
CC   (a) assessing whether a patient is afflicted with prostate cancer;
CC   (b) monitoring the progression of prostate cancer in a patient;
CC   (c) assessing the efficacy of a test compound to inhibit prostate
CC   cancer in a patient;
CC   (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC   in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound;
CC   (g) determining whether prostate cancer has metastasized in a patient;
CC   (h) assessing the aggressiveness or indolence of prostate cancer in a
CC   patient;
CC   (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ   Sequence 447 BP; 144 A; 99 C; 90 G; 114 T; 0 other;

  Query Match             80.0%;  Score 16.8;  DB 23;  Length 447;
  Best Local Similarity   90.0%;  Pred. No. 1.2e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGG 20
              |||||||||| |||| |||
Db        378 AAAGCCCCACCTTGGTGGGG 397


RESULT 15
ABA12198
ID   ABA12198 standard; cDNA; 990 BP.
XX
AC   ABA12198;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 1205.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ss.
XX
OS   Homo sapiens.
XX
PN   WO200159063-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01334.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
```

PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-541565/60.
DR P-PSDB; ABB15872.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
PS Claim 1; SEQ ID NO 1205; 1701pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 990 BP; 193 A; 336 C; 270 G; 188 T; 3 other;

Query Match 78.1%; Score 16.4; DB 22; Length 990;
Best Local Similarity 94.4%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          3 AGCCCCACCATGGTCGGG 20
              |||||||||||| |||||
Db        563 AGCCCCACCATGCTCGGG 580
```

Search completed: February 12, 2004, 04:01:27
Job time : 114.378 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 04:01:41 ; Search time 28.6829 Seconds
(without alignments)
323.156 Million cell updates/sec

Title: US-09-692-077D-14
Perfect score: 21
Sequence: 1 aaagccccaccatggtcgggt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	21	100.0	2072	4	US-09-016-434-1181	Sequence 1181, Ap
c	2	15.8	75.2	1710	4	US-09-773-426A-13	Sequence 13, Appl
	3	15.8	75.2	2754	1	US-08-196-989B-1	Sequence 1, Appli
	4	15.8	75.2	2754	2	US-08-760-936-1	Sequence 1, Appli
	5	15.8	75.2	2754	4	US-09-225-024-1	Sequence 1, Appli
c	6	15.8	75.2	2940	4	US-09-773-426A-6	Sequence 6, Appli
	7	15.4	73.3	277	4	US-09-662-250A-10	Sequence 10, Appl
	8	15.2	72.4	369	4	US-08-469-260A-28	Sequence 28, Appl
	9	15.2	72.4	369	4	US-08-488-446-28	Sequence 28, Appl
	10	15.2	72.4	369	4	US-08-467-344A-28	Sequence 28, Appl
c	11	15.2	72.4	604	4	US-09-149-476-36	Sequence 36, Appl
c	12	15.2	72.4	660	4	US-09-252-991A-4268	Sequence 4268, Ap
	13	15.2	72.4	813	4	US-09-252-991A-4484	Sequence 4484, Ap
c	14	15.2	72.4	1611	4	US-09-252-991A-4200	Sequence 4200, Ap
c	15	15.2	72.4	1911	4	US-09-050-861B-1	Sequence 1, Appli
c	16	15.2	72.4	2040	3	US-08-961-564A-1	Sequence 1, Appli
c	17	15.2	72.4	2454	1	US-07-872-673B-1	Sequence 1, Appli
c	18	15.2	72.4	6359	4	US-09-475-252-1	Sequence 1, Appli
	19	14.8	70.5	1531	1	US-08-449-986-1	Sequence 1, Appli
	20	14.8	70.5	1531	2	US-08-756-855-1	Sequence 1, Appli
	21	14.8	70.5	9723	1	US-08-083-590A-21	Sequence 21, Appl
	22	14.8	70.5	9723	3	US-08-532-384-21	Sequence 21, Appl
	23	14.6	69.5	309	4	US-09-252-991A-15583	Sequence 15583, A
c	24	14.6	69.5	417	4	US-09-328-352-462	Sequence 462, App
c	25	14.6	69.5	648	4	US-09-252-991A-1653	Sequence 1653, Ap
c	26	14.6	69.5	921	4	US-09-252-991A-13020	Sequence 13020, A
c	27	14.6	69.5	1128	2	US-08-459-101A-1	Sequence 1, Appli
c	28	14.6	69.5	1413	4	US-09-252-991A-15729	Sequence 15729, A
c	29	14.6	69.5	1527	4	US-09-252-991A-15698	Sequence 15698, A
	30	14.6	69.5	1575	4	US-09-252-991A-1521	Sequence 1521, Ap
c	31	14.6	69.5	1605	4	US-09-087-134-10	Sequence 10, Appl
	32	14.6	69.5	1791	4	US-09-252-991A-15622	Sequence 15622, A
	33	14.6	69.5	1836	4	US-09-585-174-98	Sequence 98, Appl
	34	14.6	69.5	2280	4	US-09-252-991A-12769	Sequence 12769, A
c	35	14.6	69.5	2496	1	US-08-073-384C-2	Sequence 2, Appli
c	36	14.6	69.5	2496	1	US-08-254-359A-2	Sequence 2, Appli
c	37	14.6	69.5	2496	1	US-08-483-043-2	Sequence 2, Appli
c	38	14.6	69.5	2496	1	US-08-481-238-2	Sequence 2, Appli
c	39	14.6	69.5	2496	2	US-08-471-066B-2	Sequence 2, Appli
c	40	14.6	69.5	2496	2	US-08-484-956-2	Sequence 2, Appli
c	41	14.6	69.5	2496	2	US-08-757-653-2	Sequence 2, Appli
c	42	14.6	69.5	2496	2	US-08-599-491-2	Sequence 2, Appli
c	43	14.6	69.5	2496	2	US-08-756-386-2	Sequence 2, Appli
c	44	14.6	69.5	2496	2	US-08-823-516-2	Sequence 2, Appli
c	45	14.6	69.5	2496	3	US-08-682-853A-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-016-434-1181/c
; Sequence 1181, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1181:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2072 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g178197
US-09-016-434-1181

  Query Match            100.0%;  Score 21;  DB 4;  Length 2072;
  Best Local Similarity  100.0%;  Pred. No. 0.18;
```

```
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db       1098 AAAGCCCCACCATGGTCGGGT 1078


RESULT 2
US-09-773-426A-13/c
; Sequence 13, Application US/09773426A
; Patent No. 6534302
; GENERAL INFORMATION:
;  APPLICANT: Glucksman, Maria Alexandra
;  APPLICANT:  Williamson, Mark
;  APPLICANT:  Tsia, Fong-Ying
;  APPLICANT:  Rudolph-Owen, Laura A.
;  TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302el
;  TITLE OF INVENTION:  Human Sulfatases (A CIP Application)
;  FILE REFERENCE: 35800/208398(5800-79
;  CURRENT APPLICATION NUMBER: US/09/773,426A
;  CURRENT FILING DATE:  2001-01-31
;  PRIOR APPLICATION NUMBER: US 09/495,823
;  PRIOR FILING DATE: 2000-01-31
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
;    LENGTH: 1710
;    TYPE: DNA
;    ORGANISM: homo sapiens
US-09-773-426A-13

  Query Match             75.2%;  Score 15.8;  DB 4;  Length 1710;
  Best Local Similarity   89.5%;  Pred. No. 66;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGGT 21
              || ||||||| ||||||||
Db       1070 AGACCCACCAGGGTCGGGT 1052


RESULT 3
US-08-196-989B-1
; Sequence 1, Application US/08196989B
; Patent No. 5585476
;  GENERAL INFORMATION:
;    APPLICANT:  MacLennan, A. John
;    TITLE OF INVENTION:  Molecular Cloning and Expression of
;    TITLE OF INVENTION:  G-Protein Coupled Receptors
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  US
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/196,989B
;      FILING DATE:  15-FEB-1994
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lloyd, Jeff
;      REGISTRATION NUMBER:  35,589
;      REFERENCE/DOCKET NUMBER:  MAC-100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  904-375-8100
;      TELEFAX:  904-372-5800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2754 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-196-989B-1

  Query Match             75.2%;  Score 15.8;  DB 1;  Length 2754;
  Best Local Similarity   89.5%;  Pred. No. 70;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGGT 21
              |||||||||||| ||| |
Db        139 AGCCCCACCATGGGCGGTT 157


RESULT 4
US-08-760-936-1
; Sequence 1, Application US/08760936
; Patent No. 5856443
;  GENERAL INFORMATION:
;    APPLICANT:  MacLennan, A. John
;    TITLE OF INVENTION:  Molecular Cloning and Expression of
;    TITLE OF INVENTION:  G-Protein Coupled Receptors
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  US
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/760,936
;      FILING DATE:  December 6, 1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Doran R.
;      REGISTRATION NUMBER:  38,261
;      REFERENCE/DOCKET NUMBER:  MAC-100C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  352-375-8100
;      TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2754 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-760-936-1

  Query Match             75.2%;  Score 15.8;  DB 2;  Length 2754;
  Best Local Similarity   89.5%;  Pred. No. 70;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGGT 21
              |||||||||||| ||| |
Db        139 AGCCCCACCATGGGCGGTT 157


RESULT 5
US-09-225-024-1
; Sequence 1, Application US/09225024
; Patent No. 6518414
;  GENERAL INFORMATION:
;    APPLICANT:  MacLennan, A. John
```

```
;    TITLE OF INVENTION:  Molecular Cloning and Expression of
;    TITLE OF INVENTION:  G-Protein Coupled Receptors
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik, Lloyd & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  FL
;      COUNTRY:  US
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/225,024
;      FILING DATE:  04-JAN-1999
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/760,936
;      FILING DATE:  6-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/196,989
;      FILING DATE:  15-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pace, Doran R.
;      REGISTRATION NUMBER:  38,261
;      REFERENCE/DOCKET NUMBER:  MAC-100C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  352-375-8100
;      TELEFAX:  352-372-5800
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2754 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-225-024-1

  Query Match             75.2%;  Score 15.8;  DB 4;  Length 2754;
  Best Local Similarity   89.5%;  Pred. No. 70;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGGT 21
              ||||||||||||| ||| |
Db        139 AGCCCCACCATGGGCGGTT 157


RESULT 6
US-09-773-426A-6/c
; Sequence 6, Application US/09773426A
; Patent No. 6534302
; GENERAL INFORMATION:
;  APPLICANT: Glucksman, Maria Alexandra
;  APPLICANT:  Williamson, Mark
;  APPLICANT:  Tsia, Fong-Ying
;  APPLICANT:  Rudolph-Owen, Laura A.
;  TITLE OF INVENTION: 22438, 23553, 25278, and 26212 No. 6534302el
;  TITLE OF INVENTION:  Human Sulfatases (A CIP Application)
;  FILE REFERENCE: 35800/208398(5800-79
;  CURRENT APPLICATION NUMBER: US/09/773,426A
;  CURRENT FILING DATE:  2001-01-31
;  PRIOR APPLICATION NUMBER: US 09/495,823
;  PRIOR FILING DATE: 2000-01-31
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 2940
;   TYPE: DNA
;   ORGANISM: homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (334)...(2043)
US-09-773-426A-6

  Query Match             75.2%;  Score 15.8;  DB 4;  Length 2940;
  Best Local Similarity   89.5%;  Pred. No. 70;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGGT 21
              || ||||||| ||||||||
Db       1403 AGACCCACCAGGGTCGGGT 1385


RESULT 7
US-09-662-250A-10
; Sequence 10, Application US/09662250A
; Patent No. 6368856
; GENERAL INFORMATION:
;  APPLICANT:      Brett P. Monia
;  APPLICANT:  Jacqueline Wyatt
;  TITLE OF INVENTION:      ANTISENSE MODULATION OF PHOSPHORYLASE KINASE BETA EXPRESSION
;  FILE REFERENCE:      RTS-0129
;  CURRENT APPLICATION NUMBER: US/09/662,250A
;  CURRENT FILING DATE:  2000-09-14
;  NUMBER OF SEQ ID NOS:      102
; SEQ ID NO 10
;   LENGTH: 277
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (23)...(154)
US-09-662-250A-10

  Query Match             73.3%;  Score 15.4;  DB 4;  Length 277;
  Best Local Similarity   94.1%;  Pred. No. 87;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTC 17
              ||||||||||||||| ||
Db        157 AAAGCCCCACCATGTTC 173


RESULT 8
US-08-469-260A-28
; Sequence 28, Application US/08469260A
; Patent No. 6451578
;  GENERAL INFORMATION:
;    APPLICANT:  JOHN N. SIMONS
;    APPLICANT:  TAMI J. PILOT-MATIAS
;    APPLICANT:  GEORGE J. DAWSON
;    APPLICANT:  GEORGE G. SCHLAUDER
;    APPLICANT:  SURESH M. DESAI
;    APPLICANT:  THOMAS P. LEARY
;    APPLICANT:  ANTHONY SCOTT MUERHOFF
;    APPLICANT:  JAMES C. ERKER
;    APPLICANT:  SHERI L. BUIJK
;    APPLICANT:  ISA K. MUSHAHWAR
;    TITLE OF INVENTION:  NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
;    TITLE OF INVENTION:  REAGENTS AND METHODS FOR THEIR USE
;    NUMBER OF SEQUENCES:  716
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ABBOTT LABORATORIES  D377/AP6D
;      STREET:  100 ABBOTT PARK ROAD
;      CITY:  ABBOTT PARK
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/469,260A
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/424,550
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  POREMBSKI, PRISCILLA E.
;       REGISTRATION NUMBER:  33,207
;       REFERENCE/DOCKET NUMBER:  5527.PC.01
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  708-937-6365
;       TELEFAX:  708-938-2623
;   INFORMATION FOR SEQ ID NO:  28:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  369 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-469-260A-28

  Query Match             72.4%;  Score 15.2;  DB 4;  Length 369;
  Best Local Similarity   85.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGG 20
              ||||||||||||| | | ||
Db        119 AAAGCCCCACCATAGGCCGG 138


RESULT 9
US-08-488-446-28
; Sequence 28, Application US/08488446
; Patent No. 6558898
;  GENERAL INFORMATION:
;     APPLICANT:  JOHN N. SIMONS
;     APPLICANT:  TAMI J. PILOT-MATIAS
;     APPLICANT:  GEORGE J. DAWSON
;     APPLICANT:  GEORGE G. SCHLAUDER
;     APPLICANT:  SURESH M. DESAI
;     APPLICANT:  THOMAS P. LEARY
;     APPLICANT:  ANTHONY SCOTT MUERHOFF
;     APPLICANT:  JAMES C. ERKER
;     APPLICANT:  SHERI L. BUIJK
;     APPLICANT:  ISA K. MUSHAHWAR
;     TITLE OF INVENTION:  NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
;     TITLE OF INVENTION:  REAGENTS AND METHODS FOR THEIR USE
;     NUMBER OF SEQUENCES:  716
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  ABBOTT LABORATORIES  D377/AP6D
;       STREET:  100 ABBOTT PARK ROAD
;       CITY:  ABBOTT PARK
;       STATE:  IL
;       COUNTRY:  USA
;       ZIP:  60064-3500
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/488,446
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/424,550
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  POREMBSKI, PRISCILLA E.
;       REGISTRATION NUMBER:  33,207
;       REFERENCE/DOCKET NUMBER:  5527.PC.01
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  708-937-6365
;       TELEFAX:  708-938-2623
;   INFORMATION FOR SEQ ID NO:  28:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  369 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-488-446-28

  Query Match             72.4%;  Score 15.2;  DB 4;  Length 369;
  Best Local Similarity   85.0%;  Pred. No. 1.1e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGG 20
              ||||||||||||| | | ||
Db        119 AAAGCCCCACCATAGGCCGG 138


RESULT 10
US-08-467-344A-28
; Sequence 28, Application US/08467344A
; Patent No. 6586568
;   GENERAL INFORMATION:
;       APPLICANT: JOHN N. SIMONS
;                  TAMI J. PILOT-MATIAS
;                  GEORGE J. DAWSON
;                  GEORGE G. SCHLAUDER
;                  SURESH M. DESAI
;                  THOMAS P. LEARY
;                  ANTHONY SCOTT MUERHOFF
;                  JAMES C. ERKER
;                  SHERI L. BUIJK
;                  ISA K. MUSHAHWAR
;       TITLE OF INVENTION: NON-A, NON-B. NON-C, NON-D, NON-E HEPATITIS
;                           REAGENTS AND METHODS FOR THEIR USE
;       NUMBER OF SEQUENCES: 716
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: ABBOTT LABORATORIES  D377/AP6D
;            STREET: 100 ABBOTT PARK ROAD
;            CITY: ABBOTT PARK
;            STATE: IL
;            COUNTRY: USA
;            ZIP: 60064-3500
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/08/467,344A
;            FILING DATE: 07-Jun-1995
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/424,550
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: POREMBSKI, PRISCILLA E.
;            REGISTRATION NUMBER: 33,207
;            REFERENCE/DOCKET NUMBER: 5527.PC.01
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 708-937-6365
;            TELEFAX: 708-938-2623
;   INFORMATION FOR SEQ ID NO: 28:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 369 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 28:
```

US-08-467-344A-28

Query Match 72.4%; Score 15.2; DB 4; Length 369;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 AAAGCCCCACCATGGTCGGG 20
              |||||||||||| | | ||
Db        119 AAAGCCCCACCATAGGCCGG 138
```

RESULT 11
US-09-149-476-36/c
; Sequence 36, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 72.4%; Score 15.2; DB 4; Length 604;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGG 20

Db 50 AAAGCCACTCCATGGTCTGG 31

RESULT 12
US-09-252-991A-4268/c
; Sequence 4268, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4268
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4268

Query Match 72.4%; Score 15.2; DB 4; Length 660;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGG 20

Db 575 AAAGCCCCGGCATGGCCGGG 556

RESULT 13
US-09-252-991A-4484
; Sequence 4484, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4484
; LENGTH: 813
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4484

Query Match 72.4%; Score 15.2; DB 4; Length 813;
Best Local Similarity 85.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AAAGCCCCACCATGGTCGGG 20

Db 18 AAAGCCCCGGCATGGCCGGG 37

RESULT 14
US-09-252-991A-4200/c
; Sequence 4200, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

```
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 4200
;   LENGTH: 1611
;   TYPE: DNA
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4200

  Query Match             72.4%;  Score 15.2;  DB 4;  Length 1611;
  Best Local Similarity   85.0%;  Pred. No. 1.3e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGG 20
              ||||||||  ||||| ||||
Db        699 AAAGCCCCGGCATGGCCGGG 680


RESULT 15
US-09-050-861B-1/c
; Sequence 1, Application US/09050861B
; Patent No. 6555314
; GENERAL INFORMATION:
;  APPLICANT: Payan, Donald
;  TITLE OF INVENTION: TOSO AS A TARGET FOR DRUG SCREENING
;  FILE REFERENCE: RIGL-002CON
;  CURRENT APPLICATION NUMBER: US/09/050,861B
;  CURRENT FILING DATE:  1998-03-30
;  PRIOR APPLICATION NUMBER: US/09/651,150B
;  PRIOR FILING DATE: 2000-08-30
;  PRIOR APPLICATION NUMBER: US 09/050,861
;  PRIOR FILING DATE: 1998-03-30
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1911
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-050-861B-1

  Query Match             72.4%;  Score 15.2;  DB 4;  Length 1911;
  Best Local Similarity   85.0%;  Pred. No. 1.3e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGG 20
              ||||  ||||  |||||||||
Db        860 AAAGGCCCAGGATGGTCGGG 841


Search completed: February 12, 2004, 06:07:49
Job time : 30.6829 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 04:42:03 ; Search time 119.341 Seconds
(without alignments)
648.193 Million cell updates/sec

Title: US-09-692-077D-14
Perfect score: 21
Sequence: 1 aaagccccaccatggtcgggt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	100.0	21	15	US-10-001-073-14	Sequence 14, Appl
c 2	21	100.0	65	13	US-09-908-975-4848	Sequence 4848, Ap
c 3	21	100.0	1344	9	US-09-825-923-1	Sequence 1, Appli
c 4	21	100.0	1344	15	US-10-077-870-1	Sequence 1, Appli
c 5	21	100.0	1344	15	US-10-001-073-2	Sequence 2, Appli
c 6	21	100.0	1353	9	US-09-825-923-3	Sequence 3, Appli
c 7	21	100.0	1353	15	US-10-077-870-3	Sequence 3, Appli
c 8	21	100.0	1353	15	US-10-001-073-1	Sequence 1, Appli
c 9	21	100.0	2072	12	US-10-305-720-1181	Sequence 1181, Ap
c 10	21	100.0	3274	15	US-10-225-567A-41	Sequence 41, Appl
11	16.8	80.0	4350	13	US-09-814-353-19559	Sequence 19559, A
c 12	16.4	78.1	481	11	US-09-918-995-29585	Sequence 29585, A
c 13	16.4	78.1	565	13	US-10-027-632-131352	Sequence 131352,
c 14	16.4	78.1	565	14	US-10-027-632-131352	Sequence 131352,
c 15	16.4	78.1	2262	13	US-10-220-891-38	Sequence 38, Appl
16	16.4	78.1	6997	12	US-10-094-886-137	Sequence 137, App
c 17	16.4	78.1	28438	13	US-09-820-790-3	Sequence 3, Appli
c 18	16.4	78.1	35178	13	US-10-017-161-739	Sequence 739, App
c 19	16.2	77.1	607	13	US-10-027-632-139091	Sequence 139091,
c 20	16.2	77.1	607	13	US-10-027-632-139092	Sequence 139092,
c 21	16.2	77.1	607	14	US-10-027-632-139091	Sequence 139091,
c 22	16.2	77.1	607	14	US-10-027-632-139092	Sequence 139092,
c 23	16.2	77.1	1134	15	US-10-084-817-279	Sequence 279, App
c 24	16.2	77.1	1182	13	US-10-214-519-6	Sequence 6, Appli
c 25	16.2	77.1	1182	15	US-10-071-338-6	Sequence 6, Appli
c 26	16.2	77.1	1182	15	US-10-288-985-6	Sequence 6, Appli
c 27	16.2	77.1	1182	15	US-10-289-980-6	Sequence 6, Appli
28	16.2	77.1	7193	13	US-10-214-519-1	Sequence 1, Appli
29	16.2	77.1	7193	15	US-10-071-338-1	Sequence 1, Appli
30	16.2	77.1	7193	15	US-10-288-985-1	Sequence 1, Appli
31	16.2	77.1	7193	15	US-10-289-980-1	Sequence 1, Appli
32	16	76.2	1446	12	US-10-369-493-27287	Sequence 27287, A
c 33	15.8	75.2	638	15	US-10-198-846-4384	Sequence 4384, Ap
34	15.8	75.2	716	13	US-10-027-632-25794	Sequence 25794, A
35	15.8	75.2	716	14	US-10-027-632-25794	Sequence 25794, A
c 36	15.8	75.2	1287	12	US-10-369-493-38571	Sequence 38571, A
c 37	15.8	75.2	1290	12	US-10-369-493-35369	Sequence 35369, A
c 38	15.8	75.2	1290	12	US-10-369-493-38026	Sequence 38026, A
c 39	15.8	75.2	1290	12	US-10-369-493-38181	Sequence 38181, A
c 40	15.8	75.2	1615	9	US-09-764-860-1074	Sequence 1074, Ap
c 41	15.8	75.2	1615	13	US-10-212-872-1074	Sequence 1074, Ap
c 42	15.8	75.2	1615	15	US-10-074-095-1074	Sequence 1074, Ap
c 43	15.8	75.2	1710	12	US-10-426-776-16	Sequence 16, Appl
c 44	15.8	75.2	1710	13	US-10-345-680-15	Sequence 15, Appl
c 45	15.8	75.2	1710	13	US-10-314-881-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-001-073-14
; Sequence 14, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
; APPLICANT: Liggett, Stephen
; APPLICANT: Small, Kirsten
; TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
; FILE REFERENCE: 13073-PCT
; CURRENT APPLICATION NUMBER: US/10/001,073
; CURRENT FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-001-073-14

Query Match 100.0%; Score 21; DB 15; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db          1 AAAGCCCCACCATGGTCGGGT 21
```

RESULT 2
US-09-908-975-4848/c
; Sequence 4848, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon

```
;  TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
;  TITLE OF INVENTION:  THAT POPULATE A TRANSCRIPTOME
;  FILE REFERENCE: 36688-0005
;  CURRENT APPLICATION NUMBER: US/09/908,975
;  CURRENT FILING DATE:  2001-07-20
;  PRIOR APPLICATION NUMBER: US 60/287,724
;  PRIOR FILING DATE: 2001-05-02
;  PRIOR APPLICATION NUMBER: US 60/221,607
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 32337
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4848
;   LENGTH: 65
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-09-908-975-4848

  Query Match             100.0%;  Score 21;  DB 13;  Length 65;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db         45 AAAGCCCCACCATGGTCGGGT 25


RESULT 3
US-09-825-923-1/c
; Sequence 1, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
;  APPLICANT: Snapir, Amir
;  APPLICANT:  Heinonen, Paula
;  APPLICANT:  Alhopuro, Pia
;  APPLICANT:  Karvonen, Matti
;  APPLICANT:  Koulu, Markku
;  APPLICANT:  Pesonen, Ullamari
;  APPLICANT:  Scheinin, Mika
;  APPLICANT:  Salonen, Jukka T
;  APPLICANT:  Tuomainen, Tomi-Pekka
;  APPLICANT:  Lakka, Timo A
;  APPLICANT:  Nyyss"nen, Kristiina
;  APPLICANT:  Salonen, Riitta
;  APPLICANT:  Kauhanen, Jussi
;  APPLICANT:  Valkonen, Veli-Pekka
;  TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
;  TITLE OF INVENTION:  protein, and uses thereof
;  FILE REFERENCE: Alpha-2B-AR variant
;  CURRENT APPLICATION NUMBER: US/09/825,923
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/422,985
;  PRIOR FILING DATE: 2000-05-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1341)
;   OTHER INFORMATION: Coding sequence for variant human
;   OTHER INFORMATION: alpha-2B-adrenoceptor protein
US-09-825-923-1

  Query Match             100.0%;  Score 21;  DB 9;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 4
US-10-077-870-1/c
; Sequence 1, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
;  APPLICANT: Salonen, Jukka T
;  TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
;  FILE REFERENCE: 0933-0183P
;  CURRENT APPLICATION NUMBER: US/10/077,870
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: FI 20010323
;  PRIOR FILING DATE: 2001-02-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1341)
;   OTHER INFORMATION: Coding sequence for variant human alpha-2B-adrenoceptor protein
US-10-077-870-1

  Query Match             100.0%;  Score 21;  DB 15;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 5
US-10-001-073-2/c
; Sequence 2, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
;  APPLICANT: Liggett, Stephen
;  APPLICANT:  Small, Kirsten
;  TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
;  FILE REFERENCE: 13073-PCT
;  CURRENT APPLICATION NUMBER: US/10/001,073
;  CURRENT FILING DATE:  2001-11-01
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-001-073-2

  Query Match             100.0%;  Score 21;  DB 15;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 6
US-09-825-923-3/c
; Sequence 3, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
;  APPLICANT: Snapir, Amir
;  APPLICANT:  Heinonen, Paula
;  APPLICANT:  Alhopuro, Pia
;  APPLICANT:  Karvonen, Matti
;  APPLICANT:  Koulu, Markku
;  APPLICANT:  Pesonen, Ullamari
```

```
;  APPLICANT:  Scheinin, Mika
;  APPLICANT:  Salonen, Jukka T
;  APPLICANT:  Tuomainen, Tomi-Pekka
;  APPLICANT:  Lakka, Timo A
;  APPLICANT:  Nyyss"nen, Kristiina
;  APPLICANT:  Salonen, Riitta
;  APPLICANT:  Kauhanen, Jussi
;  APPLICANT:  Valkonen, Veli-Pekka
;  TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
;  TITLE OF INVENTION:  protein, and uses thereof
;  FILE REFERENCE: Alpha-2B-AR variant
;  CURRENT APPLICATION NUMBER: US/09/825,923
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/422,985
;  PRIOR FILING DATE: 2000-05-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1350)
;   OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor
;   OTHER INFORMATION: protein
US-09-825-923-3

  Query Match             100.0%;  Score 21;  DB 9;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 7
US-10-077-870-3/c
; Sequence 3, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
;  APPLICANT: Salonen, Jukka T
;  TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
;  FILE REFERENCE: 0933-0183P
;  CURRENT APPLICATION NUMBER: US/10/077,870
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: FI 20010323
;  PRIOR FILING DATE: 2001-02-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1350)
;   OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor protein
US-10-077-870-3

  Query Match             100.0%;  Score 21;  DB 15;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 8
US-10-001-073-1/c
; Sequence 1, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
;  APPLICANT: Liggett, Stephen
;  APPLICANT:  Small, Kirsten
;  TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
;  FILE REFERENCE: 13073-PCT
;  CURRENT APPLICATION NUMBER: US/10/001,073
;  CURRENT FILING DATE:  2001-11-01
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-001-073-1

  Query Match             100.0%;  Score 21;  DB 15;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 9
US-10-305-720-1181/c
; Sequence 1181, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
;  APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
;  TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressior
;  FILE REFERENCE: PA-0002-1 CON
;  CURRENT APPLICATION NUMBER: US/10/305,720
;  CURRENT FILING DATE:  2002-11-26
;  PRIOR APPLICATION NUMBER: 09/016,434
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 1490
;  SOFTWARE: PERL Program
; SEQ ID NO 1181
;   LENGTH: 2072
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No. US20040010136A1 g178197
US-10-305-720-1181

  Query Match             100.0%;  Score 21;  DB 12;  Length 2072;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db       1098 AAAGCCCCACCATGGTCGGGT 1078


RESULT 10
US-10-225-567A-41/c
; Sequence 41, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
```

```
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
;   LENGTH: 3274
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-41

  Query Match           100.0%;  Score 21;  DB 15;  Length 3274;
  Best Local Similarity  100.0%;  Pred. No. 1.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        686 AAAGCCCCACCATGGTCGGGT 666


RESULT 11
US-09-814-353-19559
; Sequence 19559, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
;  APPLICANT: Lee, John
;  APPLICANT:  Thompson, Pamela
;  APPLICANT:  Lillie, James
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF OVARIAN CANCER
;  FILE REFERENCE: MRI-006B
;  CURRENT APPLICATION NUMBER: US/09/814,353
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: US 60/191,031
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: US 60/207,124
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 60/211,940
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: US 60/216,820
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: US 60/220,661
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: US 60/257,672
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 22037
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 19559
;   LENGTH: 4350
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 4327, 4328, 4329, 4330, 4331, 4332, 4333, 4334, 4335, 4336,
;   LOCATION: 4337, 4338, 4339, 4340, 4341, 4342, 4343, 4344, 4345, 4346,
;   LOCATION: 4347, 4348, 4349, 4350
;   OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19559

  Query Match            80.0%;  Score 16.8;  DB 13;  Length 4350;
  Best Local Similarity  90.0%;  Pred. No. 1.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGG 20
              |||||||||  ||||| ||||
Db       3482 AAAGCCCCAGCATGGCCGGG 3501


RESULT 12
US-09-918-995-29585/c
; Sequence 29585, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29585
;   LENGTH: 481
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(481)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-29585

  Query Match            78.1%;  Score 16.4;  DB 11;  Length 481;
  Best Local Similarity  94.4%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGG 20
              |||||||||||| |||||
Db        394 AGCCCCACCATGCTCGGG 377


RESULT 13
US-10-027-632-131352/c
; Sequence 131352, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 131352
;   LENGTH: 565
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-131352

  Query Match            78.1%;  Score 16.4;  DB 13;  Length 565;
  Best Local Similarity  94.4%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGG 20
              |||||||||||| |||||
Db        551 AGCCCCACCATGCTCGGG 534


RESULT 14
US-10-027-632-131352/c
; Sequence 131352, Application US/10027632
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
```

```
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 131352
;   LENGTH: 565
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-131352

  Query Match             78.1%;  Score 16.4;  DB 14;  Length 565;
  Best Local Similarity   94.4%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGG 20
              |||||||||||| |||||
Db        551 AGCCCCACCATGCTCGGG 534


RESULT 15
US-10-220-891-38/c
; Sequence 38, Application US/10220891
; Publication No. US20030207286A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWARA, AKIRA
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCES HAVING CHARACTERITICS OF ENHANCED
;  TITLE OF INVENTION:  EXPRESSION IN HUMAN NEUROBLASTOMA WITH FAVORABLE PROGNOSIS
;  TITLE OF INVENTION:  BASED ON COMPARISON BETWEEN HUMAN NEUROBLASTOMA WITH FAVORABLE
;  TITLE OF INVENTION:  PROGNOSIS AND HUMAN NEUROBLASTOMA WITH UNFAVORABLE PROGNOSIS
;  FILE REFERENCE: 7388-73435
;  CURRENT APPLICATION NUMBER: US/10/220,891
;  CURRENT FILING DATE:  2003-03-07
;  PRIOR APPLICATION NUMBER: JP 2000/140387
;  PRIOR FILING DATE: 2000-05-12
;  PRIOR APPLICATION NUMBER: JP 2000/159195
;  PRIOR FILING DATE: 2000-03-07
;  NUMBER OF SEQ ID NOS: 108
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
;   LENGTH: 2262
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-220-891-38

  Query Match             78.1%;  Score 16.4;  DB 13;  Length 2262;
  Best Local Similarity   94.4%;  Pred. No. 2.1e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 AGCCCCACCATGGTCGGG 20
              |||||||||||| |||||
Db       1487 AGCCCCACCATGCTCGGG 1470


Search completed: February 12, 2004, 06:15:44
Job time : 128.341 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 03:54:41 ; Search time 1135.79 Seconds
(without alignments)
449.373 Million cell updates/sec

Title: US-09-692-077D-14
Perfect score: 21
Sequence: 1 aaagccccaccatggtcgggt 21

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	21	100.0	1044	13	BQ880026	BQ880026 AGENCOURT
	2	18.4	87.6	261	9	AW391095	AW391095 MR3-ST020
	3	17.8	84.8	269	9	AI582395	AI582395 tr97b09.x
c	4	17.8	84.8	376	9	AU166948	AU166948 AU166948
	5	17.8	84.8	724	9	AI419320	AI419320 tf27e11.x
	6	17.8	84.8	921	9	AL566688	AL566688 AL566688
c	7	17.8	84.8	956	13	BQ923897	BQ923897 AGENCOURT
c	8	17.4	82.9	212	9	AI216547	AI216547 qm36h04.x
c	9	17.4	82.9	487	9	AI869725	AI869725 wl98h12.x
	10	17.4	82.9	740	28	BH794919	BH794919 ME_MBa000
	11	17.4	82.9	1485	12	BI687242	BI687242 603314820
c	12	17	81.0	696	28	AZ571445	AZ571445 287PvG04
c	13	17	81.0	703	12	BP111519	BP111519 BP111519
c	14	17	81.0	960	12	BM469354	BM469354 AGENCOURT
c	15	17	81.0	1017	12	BQ053165	BQ053165 AGENCOURT
c	16	16.8	80.0	376	13	BY401935	BY401935 BY401935
	17	16.8	80.0	407	13	BX316223	BX316223 BX316223
	18	16.8	80.0	449	9	AI494022	AI494022 qz97h04.x
c	19	16.8	80.0	492	14	CA912228	CA912228 PCSC21166
	20	16.8	80.0	659	9	AW161813	AW161813 au70g04.x
	21	16.8	80.0	697	14	CB529586	CB529586 UI-H-FT2-
	22	16.8	80.0	726	28	BH794448	BH794448 ME_MBa000
	23	16.8	80.0	894	29	CNS031RL	AL223914 Tetraodon
c	24	16.8	80.0	896	13	BU603088	BU603088 AGENCOURT
c	25	16.8	80.0	912	10	BF307809	BF307809 601890896
c	26	16.8	80.0	946	9	AL549851	AL549851 AL549851
	27	16.8	80.0	950	9	AL533975	AL533975 AL533975
	28	16.8	80.0	957	29	BZ551182	BZ551182 pacs1-60_
c	29	16.8	80.0	1029	12	BQ061158	BQ061158 AGENCOURT
	30	16.8	80.0	1059	13	BU587459	BU587459 AGENCOURT
	31	16.8	80.0	1146	10	BG242458	BG242458 602354253
	32	16.4	78.1	212	10	BF946030	BF946030 CM0-NN115
	33	16.4	78.1	333	10	BF811542	BF811542 CM2-CI017
	34	16.4	78.1	333	13	BY149362	BY149362 BY149362
	35	16.4	78.1	341	13	BY195012	BY195012 BY195012
	36	16.4	78.1	345	13	BY214037	BY214037 BY214037
	37	16.4	78.1	354	13	BY337375	BY337375 BY337375
	38	16.4	78.1	360	13	BY200267	BY200267 BY200267
c	39	16.4	78.1	365	14	CB069817	CB069817 is29e01.y
c	40	16.4	78.1	379	9	AA448022	AA448022 zw97b03.r
	41	16.4	78.1	430	10	BE517145	BE517145 WHE0624_F
c	42	16.4	78.1	433	9	AA448451	AA448451 zw96h03.r
	43	16.4	78.1	442	9	AA447622	AA447622 zw97b03.s
c	44	16.4	78.1	446	12	BI713098	BI713098 ie02b10.y
	45	16.4	78.1	448	9	AA448314	AA448314 zw96h03.s

ALIGNMENTS

RESULT 1
BQ880026/c
LOCUS BQ880026 1044 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT_8113358 Lupski_dorsal_root_ganglion Homo sapiens cDNA clone IMAGE:6179035 5', mRNA sequence.
ACCESSION BQ880026
VERSION BQ880026.1 GI:22272034
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1044)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13559 row: i column: 20
High quality sequence stop: 430.

```
FEATURES             Location/Qualifiers
     source          1. .1044
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6179035"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      182 a    334 c    345 g    183 t
ORIGIN

  Query Match             100.0%;  Score 21;  DB 13;  Length 1044;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||||||||||||||||
Db        223 AAAGCCCCACCATGGTCGGGT 203


RESULT 2
AW391095
LOCUS       AW391095                261 bp    mRNA    linear   EST 04-FEB-2000
DEFINITION  MR3-ST0203-131299-111-f11 ST0203 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW391095
VERSION     AW391095.1  GI:6895754
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 261)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-ST0203-
            131299-111-f11&t3=1999-12-13&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 10
            High quality sequence stop: 63.
FEATURES             Location/Qualifiers
     source          1. .261
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="ST0203"
                     /note="Organ: stomach; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT       67 a     54 c     80 g     60 t
ORIGIN

  Query Match             87.6%;  Score 18.4;  DB 9;  Length 261;
  Best Local Similarity   95.0%;  Pred. No. 3.6e+02;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAGCCCCACCATGGTCGGGT 21
              |||||||||||||||| ||||
Db         54 AAGCCCCACCATGGTTGGGT 73


RESULT 3
AI582395
LOCUS       AI582395                269 bp    mRNA    linear   EST 14-DEC-1999
DEFINITION  tr97b09.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2227001 3',
            mRNA sequence.
ACCESSION   AI582395
VERSION     AI582395.1  GI:4568292
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 269)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11548-013
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1875   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 141
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .269
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2227001"
                     /tissue_type="adenocarcinoma"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Pan1"
                     /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.72 kb. Life Technologies catalog #:
                     11548-013"
BASE COUNT       80 a     44 c     49 g     96 t
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 269;
  Best Local Similarity   90.5%;  Pred. No. 6.9e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              | |||||||  ||||||||||
Db        209 AGAGCCCCAACATGGTCGGGT 229


RESULT 4
AU166948/c
LOCUS       AU166948                376 bp    mRNA    linear   EST 03-APR-2002
DEFINITION  AU166948 Rice callus (2001) Oryza sativa (japonica cultivar-group)
```

```
            cDNA clone C60162, mRNA sequence.
ACCESSION   AU166948
VERSION     AU166948.1  GI:12405347
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from callus (2001)
  JOURNAL   Unpublished
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'.
FEATURES             Location/Qualifiers
     source          1. .376
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="C60162"
                     /tissue_type="callus"
                     /clone_lib="Rice callus (2001)"
BASE COUNT      102 a     81 c     81 g    106 t      6 others
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 376;
  Best Local Similarity   90.5%;  Pred. No. 7.2e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              || |||||||||||||| ||||
Db        132 AATGCCCCACCATGGTAGGGT 112


RESULT 5
AI419320
LOCUS       AI419320                724 bp    mRNA    linear   EST 30-MAR-1999
DEFINITION  tf27e11.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2097452 3'
            similar to gb:X52255_rna1 CYSTATIN C PRECURSOR (HUMAN);, mRNA
            sequence.
ACCESSION   AI419320
VERSION     AI419320.1  GI:4265251
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 724)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 835    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 412.
FEATURES             Location/Qualifiers
     source          1. .724
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2097452"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Brn23"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      141 a    204 c    242 g    134 t      3 others
ORIGIN

  Query Match             84.8%;  Score 17.8;  DB 9;  Length 724;
  Best Local Similarity   90.5%;  Pred. No. 8e+02;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGGGT 21
              |||||||| |||||||||| |
Db        682 AAAGCCCCGCCATGGTCGGTT 702


RESULT 6
AL566688
LOCUS       AL566688                921 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL566688 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF019YH12 3-PRIME, mRNA sequence.
ACCESSION   AL566688
VERSION     AL566688.2  GI:31289656
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 921)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12919302.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 8536.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DF019DD06NP1&cluster=8536.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DF019DD06NP1.
FEATURES             Location/Qualifiers
     source          1. .921
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DF019YH12"
                     /tissue_type="FETAL BRAIN"
                     /dev_stage="fetal"
                     /clone_lib="Homo sapiens FETAL BRAIN"
                     /note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-strand cDNA was digested with Not I and
```

cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT 231 a 184 c 269 g 176 t 61 others
ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 921;
Best Local Similarity 90.5%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 AAAGCCCCACCATGGTCGGGT 21
              ||||||||||||  ||| ||||
Db        544 AAAGCCCCACCAGGGTGGGGT 564
```

RESULT 7
BQ923897/c
LOCUS BQ923897 956 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT_8854334 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6473613 5', mRNA sequence.
ACCESSION BQ923897
VERSION BQ923897.1 GI:22338928
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 956)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14009 row: g column: 22
High quality sequence stop: 413.
FEATURES Location/Qualifiers
source 1. .956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6473613"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. "
BASE COUNT 247 a 248 c 228 g 233 t
ORIGIN

Query Match 84.8%; Score 17.8; DB 13; Length 956;
Best Local Similarity 90.5%; Pred. No. 8.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 AAAGCCCCACCATGGTCGGGT 21
              ||| ||||||||||||  ||||
Db        936 AAAACCCCACCATGGTGGGGT 916
```

RESULT 8
AI216547/c
LOCUS AI216547 212 bp mRNA linear EST 02-DEC-1998
DEFINITION qm36h04.x1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1883959 3', mRNA sequence.
ACCESSION AI216547
VERSION AI216547.1 GI:3785588
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 212)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 589 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 177.
FEATURES Location/Qualifiers
source 1. .212
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1883959"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu5"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 53 a 51 c 60 g 48 t
ORIGIN

Query Match 82.9%; Score 17.4; DB 9; Length 212;
Best Local Similarity 94.7%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 AAAGCCCCACCATGGTCGG 19
              ||||||||||||||||| ||
Db        104 AAAGCCCCACCATGGTTGG 86
```

RESULT 9
AI869725/c
LOCUS AI869725 487 bp mRNA linear EST 07-MAR-2000
DEFINITION wl98h12.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2432999 3' similar to gb:L11285 DUAL SPECIFICITY MITOGEN-ACTIVATED PROTEIN KINASE KINASE 2 (HUMAN);, mRNA sequence.
ACCESSION AI869725
VERSION AI869725.1 GI:5543693
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 487)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,

```
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1703    Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 364.
FEATURES              Location/Qualifiers
     source           1. .487
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2432999"
                      /tissue_type="anaplastic oligodendroglioma"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Brn25"
                      /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer (5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      Library is normalized, and was constructed by Bento
                      Soares and M.Fatima Bonaldo."
BASE COUNT      103 a    160 c    160 g     64 t
ORIGIN

  Query Match             82.9%;  Score 17.4;  DB 9;  Length 487;
  Best Local Similarity   94.7%;  Pred. No. 1.1e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGG 19
              |||||||| ||||||||||
Db        479 AAAGCCCCCCCATGGTCGG 461


RESULT 10
BH794919
LOCUS       BH794919                740 bp    DNA     linear   GSS 02-APR-2002
DEFINITION  ME_MBa0003L01r Manihot esculenta Manihot esculenta genomic clone
            ME_MBa0003L01r, genomic survey sequence.
ACCESSION   BH794919
VERSION     BH794919.1  GI:19893527
KEYWORDS    GSS.
SOURCE      Manihot esculenta (cassava)
  ORGANISM  Manihot esculenta
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Malpighiales; Euphorbiaceae; Manihot.
REFERENCE   1  (bases 1 to 740)
  AUTHORS   Tomkins,J.P., Fregene,M., Main,D., Goicoechea,J.L., Blackmon,B.,
            Atkins,M., Tohme,J. and Wing,R.A.
  TITLE     New Genomic Resources for Cassava (Manihot esculenta): Development
            of a Deep-Coverage BAC Library and Preliminary STC Analysis
  JOURNAL   Unpublished
COMMENT     Contact: Tomkins J
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson University, Clemson, SC 29634, USA
            Tel: 864 656 6419
            Fax: 864 656 4293
            Email: jtmkns@clemson.edu
            Total High Quality bases = 176
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence start: 38
            High quality sequence stop: 514.
FEATURES              Location/Qualifiers
     source           1. .740
                      /organism="Manihot esculenta"
                      /mol_type="genomic DNA"
                      /strain="MECW72"
                      /db_xref="taxon:3983"
                      /clone="ME_MBa0003L01r"
                      /tissue_type="Leaf"
                      /lab_host="E. coli"
                      /clone_lib="Manihot esculenta"
                      /note="Vector: pCUGIBAC-1; Site_1: HindIII; Site_2: NotI;
                      For more details on library preparation and sequence
                      analysis see
                      http://www.genome.clemson.edu/projects/stc/cassava/ME_MBa
                      To order clones from this library see
                      http://www.genome.clemson.edu/orders "
BASE COUNT      174 a    216 c    158 g    191 t      1 others
ORIGIN

  Query Match             82.9%;  Score 17.4;  DB 28;  Length 740;
  Best Local Similarity   94.7%;  Pred. No. 1.2e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGG 19
              ||| ||||||||||||||||
Db        227 AAAACCCCACCATGGTCGG 245


RESULT 11
BI687242
LOCUS       BI687242                1485 bp    mRNA    linear   EST 18-SEP-2001
DEFINITION  603314820F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5354908 5',
            mRNA sequence.
ACCESSION   BI687242
VERSION     BI687242.1  GI:15649870
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1485)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11902  row: g  column: 05.
FEATURES              Location/Qualifiers
     source           1. .1485
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="FVB/N"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:5354908"
                      /sex="female, virgin"
                      /tissue_type="infiltrating ductal carcinoma"
                      /dev_stage="5 months"
                      /lab_host="DH10B"
                      /clone_lib="NCI_CGAP_Mam6"
                      /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Library constructed by Life Technologies. Investigator
                      providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      507 a    342 c    327 g    306 t      3 others
ORIGIN

  Query Match             82.9%;  Score 17.4;  DB 12;  Length 1485;
```

```
  Best Local Similarity   94.7%;  Pred. No. 1.4e+03;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTCGG 19
              ||||||||||||||||| |||
Db       1404 AAAGCCCCACCATGGGCGG 1422


RESULT 12
AZ571445/c
LOCUS       AZ571445                696 bp    DNA     linear   GSS 15-MAY-2001
DEFINITION  287PvG04 Pv MBN #30 Plasmodium vivax genomic 3', genomic survey
            sequence.
ACCESSION   AZ571445
VERSION     AZ571445.1  GI:13983537
KEYWORDS    GSS.
SOURCE      Plasmodium vivax (malaria parasite P. vivax)
  ORGANISM  Plasmodium vivax
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 696)
  AUTHORS   Carlton,J.M.-R. and Dame,J.B.
  TITLE     The Plasmodium vivax and P. berghei gene sequence tag projects
  JOURNAL   Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
COMMENT     Contact: Dame JB
            Dept. of Pathobiology, College of Veterinary Medicine
            University of Florida
            2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
            Tel: 352 392 4700
            Fax: 352 392 9704
            Email: damej@mail.vetmed.ufl.edu
            Seq primer: M13(-20) forward
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1. .696
                     /organism="Plasmodium vivax"
                     /mol_type="genomic DNA"
                     /strain="Salvador I (Collins, W. 1972. J. Parasitol. 69,
                     497-598)"
                     /db_xref="taxon:5855"
                     /dev_stage="asexual blood forms"
                     /lab_host="Saimiri boliviensis"
                     /clone_lib="Pv MBN #30"
                     /note="Vector: pBluescript SK(+) vector DNA, phagemid
                     excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V;
                     Host leukocytes were extracted from P. vivax infected
                     blood using the following methods: first, infected blood
                     was activated by the addition of 0.5 ml of ADP (40mg/ml)
                     per 10 ml blood. Then blood was passed over a column of
                     acid washed 0.1 mm glass beads, then through a Plasmodipur
                     filter, followed by passage through a column of pre-wet
                     Whatman CF11 powder (1:2 ratio volume of blood to CF11),
                     and finally centrifuged through a 50% Percoll density
                     cushion. Purified DNA was digested with mung bean nuclease
                     in the presence of 44% formamide at 50oC as described
                     (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.
                     Nucleic Acids Research 16:6883-6896). Digested DNA was
                     blunt-ended using T4 DNA Polymerase and size fractionated
                     over a Sepharose CL-2B column. Fractions in the size range
                     500bp-4kb were ligated into the Eco RV site of pBluescript
                     SK(+), and E. coli XL-10 Gold transformed with the
                     ligation mixture."
BASE COUNT      181 a    198 c    185 g    126 t      6 others
ORIGIN

  Query Match             81.0%;  Score 17;  DB 28;  Length 696;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTC 17
              |||||||||||||||||
Db        696 AAAGCCCCACCATGGTC 680


RESULT 13
BP111519/c
LOCUS       BP111519                703 bp    mRNA    linear   EST 11-FEB-2003
DEFINITION  BP111519 ORCS bovine utero-placenta cDNA Bos taurus cDNA clone
            ORCS11942 5', mRNA sequence.
ACCESSION   BP111519
VERSION     BP111519.1  GI:28313809
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 703)
  AUTHORS   Ishiwata,H., Katsuma,S., Kizaki,K., Patel,O.V., Nakano,H.,
            Takahashi,T., Imai,K., Hirasawa,A., Shiojima,S., Ikawa,H., Suzuki
            ,Y., Tsujimoto,G., Izaike,Y., Todoroki,J. and Hashizume,K.
  TITLE     Characterization of gene expression profiles in early bovine
            pregnancy using a custom cDNA microarray
  JOURNAL   Mol. Reprod. Dev. 65 (1), 9-18 (2003)
  MEDLINE   22544902
   PUBMED   12658628
COMMENT     Contact: Gozoh Tsujimoto
            Department of Molecular, Cell Pharmacology
            National Research Institute for Child Health and Development
            3-35-31 Taishido, Setagaya, Tokyo 154-8567, Japan
            Tel: 81-3-3149-2476
            Fax: 81-3-3149-1252
            Email: gtsujimoto@nch.go.jp
            This work was performed to collaborate with Developmental Biology
            Department, National Institute of Agrobiological Sciences. Address:
            2 Ikenodai, Tsukuba, Ibaraki, 305-8602 Japan. Phone & Fax:
            81-29-838-8633 e-mail: kazuha@affrc.go.jp
            This work was funded by Organized Research Combination System (ORCS
            ) project of Ministry of Education, Culture, Sports, Science and
            Technology.
FEATURES             Location/Qualifiers
     source          1. .703
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /clone="ORCS11942"
                     /tissue_type="mixture of uterus and placenta"
                     /dev_stage="adult"
                     /clone_lib="ORCS bovine utero-placenta cDNA"
BASE COUNT      163 a    152 c    203 g    183 t      2 others
ORIGIN

  Query Match             81.0%;  Score 17;  DB 12;  Length 703;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTC 17
              |||||||||||||||||
Db        166 AAAGCCCCACCATGGTC 150


RESULT 14
BM469354/c
LOCUS       BM469354                960 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6480665 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5551482
            5', mRNA sequence.
ACCESSION   BM469354
VERSION     BM469354.1  GI:18518396
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 960)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12265  row: e  column: 19
            High quality sequence stop: 545.
FEATURES              Location/Qualifiers
     source           1. .960
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5551482"
                      /tissue_type="lymphoma, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_85"
                      /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                      Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                      Average insert size 1.867 kb. Library enriched for
                      full-length clones and constructed by Life Technologies.
                      Note: this is a NIH_MGC Library."
BASE COUNT      250 a    221 c    248 g    241 t
ORIGIN

  Query Match             81.0%;  Score 17;  DB 12;  Length 960;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTC 17
              |||||||||||||||||
Db        810 AAAGCCCCACCATGGTC 794


RESULT 15
BQ053165/c
LOCUS       BQ053165                1017 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6821899 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5934766
            5', mRNA sequence.
ACCESSION   BQ053165
VERSION     BQ053165.1  GI:19812505
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1017)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2121  row: g  column: 23
            High quality sequence stop: 825.
FEATURES              Location/Qualifiers
     source           1. .1017
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5934766"
                      /tissue_type="natural killer cells, cell line"
                      /lab_host="DH10B (phage-resistant)"
                      /clone_lib="NIH_MGC_106"
                      /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                      into EcoRI/XhoI sites using the following 5' adaptor:
                      GGCACGAG(G). Library constructed  by Ling Hong in the
                      laboratory of Gerald M. Rubin (University of  California,
                      Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                      Superscript II RT (Life Technologies). Note: this is a
                      NIH_MGC Library."
BASE COUNT      279 a    228 c    264 g    245 t      1 others
ORIGIN

  Query Match             81.0%;  Score 17;  DB 12;  Length 1017;
  Best Local Similarity   100.0%;  Pred. No. 2e+03;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAGCCCCACCATGGTC 17
              |||||||||||||||||
Db       1004 AAAGCCCCACCATGGTC 988


Search completed: February 12, 2004, 06:05:52
Job time : 1142.79 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 01:54:29 ; Search time 585.854 Seconds
(without alignments)
1396.582 Million cell updates/sec

Title: US-09-692-077D-15
Perfect score: 20
Sequence: 1 ctgatcgccaaacgagcaac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	6	AX350503	AX350503 Sequence
2		17.4	87.0	45260	3	AC134175	AC134175 Drosophil
3		17.4	87.0	154555	8	AC135958	AC135958 Oryza sat
4		17	85.0	176534	2	AC135242	AC135242 Oryza sat
5		16.8	84.0	1541	3	AY058358	AY058358 Drosophil
6		16.8	84.0	71284	8	F21B23	AF262038 Arabidops
7	c	16.8	84.0	91268	8	AC006837	AC006837 Arabidops
8	c	16.8	84.0	91436	8	NCB14A6	AL670007 Neurospor
9	c	16.8	84.0	96232	8	ATAC010927	AC010927 Arabidops
10		16.8	84.0	173613	3	AC007475	AC007475 Drosophil
11		16.8	84.0	192763	3	AC007474	AC007474 Drosophil
12	c	16.8	84.0	194634	2	AC020286	AC020286 Drosophil
13		16.8	84.0	206732	2	AC109049	AC109049 Rattus no
14		16.8	84.0	259345	2	AC102969	AC102969 Rattus no
15		16.8	84.0	261157	2	AC103330	AC103330 Rattus no
16		16.8	84.0	284749	2	AC113251	AC113251 Rattus no
17		16.8	84.0	292919	3	AE003823	AE003823 Drosophil
18	c	16.8	84.0	302101	1	AE016784	AE016784 Pseudomon
19	c	16.8	84.0	324050	1	BX251410	BX251410 Tropherym
20		16.8	84.0	324227	1	AE016852	AE016852 Tropherym
21		16.4	82.0	706	9	HSA332772	AJ332772 Homo sapi
22		16.4	82.0	926	7	P8080COR	D00360 Bacteriopha
23	c	16.4	82.0	97631	2	AC141670	AC141670 Apis mell
24	c	16.4	82.0	152224	2	AC141680	AC141680 Apis mell
25	c	16.4	82.0	217013	2	AC106439	AC106439 Rattus no
26		16	80.0	184862	9	AC093905	AC093905 Homo sapi
27		16	80.0	301700	1	AP005948	AP005948 Bradyrhiz
28		15.8	79.0	1443	1	AF351223	AF351223 Unculture
29	c	15.8	79.0	2000	6	AX656134	AX656134 Sequence
30		15.8	79.0	4269	8	SCYKL210W	Z28210 S.cerevisia
31		15.8	79.0	4795	8	SCUBA1G	X55386 Yeast UBA1
32	c	15.8	79.0	7799	10	AB053223S1	AB053223 Mus muscu
33		15.8	79.0	18813	8	T9E19	AF104920 Arabidops
34	c	15.8	79.0	41338	2	AC017942	AC017942 Drosophil
35		15.8	79.0	41625	3	AF022973	AF022973 Caenorhab
36		15.8	79.0	48404	8	AC079280	AC079280 Arabidops
37		15.8	79.0	49830	6	AX059541	AX059541 Sequence
38		15.8	79.0	51315	9	AC134508	AC134508 Homo sapi
39	c	15.8	79.0	57303	2	AC020240	AC020240 Drosophil
40	c	15.8	79.0	61849	3	AC004306	AC004306 Drosophil
41		15.8	79.0	63989	8	AB013393	AB013393 Arabidops
42	c	15.8	79.0	70098	8	AB025605	AB025605 Arabidops
43	c	15.8	79.0	73663	8	AP002029	AP002029 Arabidops
44	c	15.8	79.0	80393	8	AP000382	AP000382 Arabidops
45		15.8	79.0	81414	8	AB024037	AB024037 Arabidops

ALIGNMENTS

RESULT 1
AX350503
LOCUS AX350503 20 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 15 from Patent WO0179561.
ACCESSION AX350503
VERSION AX350503.1 GI:18616099
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liggett,S.B. and Small,K.M.
TITLE Alpha-2 adrenergic receptor polymorphisms
JOURNAL Patent: WO 0179561-A 15 25-OCT-2001;

Liggett, Stephen B. (US) ; Small, Kersten M. (US)
FEATURES Location/Qualifiers
source 1. .20
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 7 a 7 c 4 g 2 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGATCGCCAAACGAGCAAC 20
||||||||||||||||||||
Db 1 CTGATCGCCAAACGAGCAAC 20

RESULT 2
AC134175
LOCUS AC134175 45260 bp DNA linear INV 24-SEP-2002
DEFINITION Drosophila pseudoobscura FOSMID DPSF1-548A9 (Children's Hospital Oakland Research Institute Drosophila pseudoobscura FOSMID Library) complete sequence.
ACCESSION AC134175
VERSION AC134175.1 GI:23306009
KEYWORDS HTG.
SOURCE Drosophila pseudoobscura
ORGANISM Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 45260)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Adams,C., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Baca,E., Bandaranaike,D., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya Khurana, Draper,H., D'Souza,L., Dugan-Rocha,S., Dunn,A., Durbin,K., Eaves,K., Egan,A., Escotto,M., Eugene,C., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Lopez,J., Lorensuhewa,L., Loulseged,H., Lozado,R., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., Mcleod,M., Mcneill,T., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Reigh,R., Ren,Y., Reuter,M., Richards,S., Rives,C., Rodkey,T., Rojas,A., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shen,H., Sisson,I., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Weinstock,G. and Gibbs,R.A.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 45260)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (24-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.
FEATURES Location/Qualifiers
source 1. .45260
/organism="Drosophila pseudoobscura"
/mol_type="genomic DNA"
/db_xref="taxon:7237"
/clone="DPSF1-548A9"
BASE COUNT 11756 a 10970 c 10929 g 11605 t
ORIGIN

Query Match 87.0%; Score 17.4; DB 3; Length 45260;
Best Local Similarity 94.7%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGATCGCCAAACGAGCAA 19
||||| ||||||||||||
Db 30040 CTGATTGCCAAACGAGCAA 30058

RESULT 3
AC135958
LOCUS AC135958 154555 bp DNA linear PLN 25-APR-2003
DEFINITION Oryza sativa chromosome 3 BAC OSJNBa0059E14 genomic sequence, complete sequence.
ACCESSION AC135958
VERSION AC135958.2 GI:29837774
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 154555)
AUTHORS Buell,C.R., Yuan,Q., Ouyang,S., Liu,J., Gansberger,K., Jones,K.M., Overton II,L.L., Tsitrin,T., Kim,M.M., Bera,J.J., Jin,S.S., Fadrosh,D.W., Tallon,L.J., Koo,H., Zismann,V., Hsiao,J., Blunt,S., Vanaken,S.S., Riedmuller,S.B., Utterback,T.T., Feldblyum,T.V., Yang,Q.Q., Haas,B.J., Suh,B.B., Peterson,J.J., Quackenbush,J., White,O., Salzberg,S.L. and Fraser,C.M.
TITLE Oryza sativa chromosome 3 BAC OSJNBa0059E14 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 154555)
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (01-NOV-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 3 (bases 1 to 154555)
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (15-APR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REFERENCE 4 (bases 1 to 154555)
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
COMMENT On Apr 15, 2003 this sequence version replaced gi:24462343.
Address all correspondence to:rice@tigr.org

BAC clone OSJNBa0059E14 is from Oryza sativa chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC

```
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including Fgenesh (http://www.softberry.com/),
genscan and Genscan+ (Chris Burge,
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHMM (Mark Borodovsky,
http://genemark.biology.gatech.edu/GeneMark/), and GeneSplicer
(Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as  unknown  proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as  hypothetical  proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
identified by repeatmasker (Arian Smit,
http://ftp.genome.washington.edu/RM/RepeatMasker.html).

This BAC overlaps with rice BAC OSJNBb0015I02 (AC135563).
FEATURES             Location/Qualifiers
     source          1. .154555
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="genomic DNA"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /chromosome="3"
                     /map="near C393A"
                     /clone="OSJNBa0059E14"
                     /note="japonica cultivar-group"
     repeat_region   1373. .1396
                     /rpt_family="AT_rich"
     gene            complement(2338. .3123)
                     /gene="OSJNBa0059E14.1"
                     /note="EST AU165836,AU182644 from this gene"
     mRNA            complement(join(<2338. .2570,2665. .>3123))
                     /gene="OSJNBa0059E14.1"
     CDS             complement(join(2499. .2570,2665. .3123))
                     /gene="OSJNBa0059E14.1"
                     /codon_start=1
                     /product="unknown protein"
                     /protein_id="AAP21417.1"
                     /db_xref="GI:30103004"
                     /translation="MAAIANLELSTHTHTHTQRVCSIDRPSPRAHDGERSSQPRRGGG
                     GDRGVCGCRGGGGGRRGRGGGVVRRRGERAGAVRAVPAGRRGAAGRSVLRRREGAARD
                     GGHGGGAARAVQVPGAVRAVVRRAPRPRAAPPGALQARTRHPRRRRHRLQQDILSGSG
                     GDDDAATWREFVAIGA"
     repeat_region   2782. .3013
                     /rpt_family="(CCGCG)n"
     gene            4169. .5246
                     /gene="OSJNBa0059E14.2"
                     /note="similar to ubiquitin GB:BAC06474  GI:22004048
                     (Ciona savignyi)"
     mRNA            join(<4169. .4410,4495. .4975,5067. .>5246)
                     /gene="OSJNBa0059E14.2"
     CDS             join(4169. .4410,4495. .4975,5067. .5246)
                     /gene="OSJNBa0059E14.2"
                     /codon_start=1
                     /product="putative ubiquitin"
                     /protein_id="AAP21418.1"
                     /db_xref="GI:30103005"
                     /translation="MRIMVRTLRGDRVALDVDGATTTVAQVKGMVMARERIAVAMQRL
                     FFAGRCLDDDHRTLADYGVRHDSVVFLSLRLATDAYQTEMHNVRLMQPETATAKQEMH
                     QQQQQQLHVHVAADDEEKAIKRKPVSRRALRKILSRLQVDAWTSQHDAKFLDLLLRHT
                     GGGGGARNVGELTGEDWSSIRAELNAATGSGFPVEELQRRLGEFRREFEAASRIKNHP
                     RFSYDPRRRVVVAKQADWKNYILENPEAAAYEGRSPRHLGRLRAIFSGDGGGGGGGGA
                     KCRETKARSCLRKLLRNFRLRFKL"
     repeat_region   5138. .5161
                     /rpt_family="(CGG)n"
     gene            9301. .11079
                     /gene="OSJNBa0059E14.3"
                     /note="contains Pfam profile PF03763 (remorin, C terminal
                     region); EST BE041106,D41920,AU097433 from this gene"
     mRNA            join(9301. .10065,10558. .11079)
                     /gene="OSJNBa0059E14.3"
     CDS             join(9403. .10065,10558. .10749)
                     /gene="OSJNBa0059E14.3"
                     /codon_start=1
                     /product="remorin-like protein"
                     /protein_id="AAP21420.1"
                     /db_xref="GI:30103007"
                     /translation="MLHEQHAPPPQPEPEVSLQLSAPATAADDVAAGDDEEVTVVTTY
                     RDIHPLTPPSPTTTTTPPTRLGSAAYSWDTASSHRSVSSEEQFMTMSREFTAMVAAGT
                     TMQTGPNDGNNGGDQLTSIGEDELEETNPLAIVPDSHPIATPARSRASQLEVVPAAGP
                     SPAPPVEARQVKKEEVETKVSAWQTAEVAKINNRFKREEVVINGWETEQVEKASAWLK
                     KIERKLDEQRAKALERTQNDIAKARRKAEEKRASAEAKRGLKLAKVLELANFMKAVGR
                     VPTKRSFF"
     repeat_region   9422. .9587
                     /rpt_family="(CCG)n"
     gene            complement(12645. .14126)
                     /gene="OSJNBa0059E14.4"
                     /note="similar to immediate early salicylate induced
                     glucosyltransferase GB:AAB36653  GI:1685005 (Nicotiana
                     tabacum)"
     mRNA            complement(<12645. .>14126)
                     /gene="OSJNBa0059E14.4"
     CDS             complement(12645. .14126)
                     /gene="OSJNBa0059E14.4"
                     /codon_start=1
                     /product="putative immediate-early salicylate-induced
                     glucosyltransferase"
                     /protein_id="AAP21423.1"
                     /db_xref="GI:30103010"
                     /translation="MAAATADGHGGRRRLRVFFLPFFARGHLIPMTDLACLMAAASTD
                     AVEVEATMAVTPANAAAIAATVAGNAAVRVVCYPFPDVGLARGVECLGAAAAHDTWRV
                     YRAVDLSRPAHESLLRHHRPDAIVADVPFWWATGVAAELGVPRLTFNPVGVFPQLAMN
                     NLVAVRPDIVRGGADGPPVTVPGMPGGREITIPVSELPDFLVQDDHLSMSWDRIKASQ
                     LAGFGVVVNTFAALEAPYCDEFSRVDARRAYFVGPVSQPSRAAAAAVRRGGDGDVDCL
                     RWLSTKPSQSVVYVCFGSWAHFSVTQTRELALGLEASNQPFLWVIRSDSGDGGGERWE
                     PEGWERRMEGRGMVVRGWAPQLAVLAHPSVGAFVTHCGWNSVLEAAAAGVPALTWPLV
                     FEQFINERLVTEVAAFGARVWEDGGGKRGVRAREAETVPAGVIARAVAGFMAGGGGRR
                     ERAAAMATALAESARVAVGENGSSWRDIRRLIQDLTDATASQP"
     repeat_region   13319. .13354
                     /rpt_family="GC_rich"
     repeat_region   13913. .13963
                     /rpt_family="(CGG)n"
     gene            15663. .21236
                     /gene="OSJNBa0059E14.5"
                     /note="highly similar to cellulose synthase catalytic
                     subunit  GB:AAC39336  GI:2827143 (Arabidopsis thaliana);
                     EST AU064228,AU176398, CB096725,AA751514 from this gene"
     mRNA            join(15663. .16059,16193. .16379,16712. .16962,17065. .17164,
                     17243. .17429,17725. .17991,18085. .18430,18519. .18656,
                     18739. .18864,18950. .19162,19322. .19577,19773. .19975,
                     20219. .20569,20652. .>21236)
                     /gene="OSJNBa0059E14.5"
     CDS             join(16048. .16059,16193. .16379,16712. .16962,17065. .17164,
                     17243. .17429,17725. .17991,18085. .18430,18519. .18656,
                     18739. .18864,18950. .19162,19322. .19577,19773. .19975,
                     20219. .20569,20652. .21236)
                     /gene="OSJNBa0059E14.5"
                     /codon_start=1
                     /product="putative cellulose synthase catalytic subunit"
                     /protein_id="AAP21426.1"
                     /db_xref="GI:30103013"
                     /translation="MDGAKSGKQCHVCQICGDGVGTAADGELFTACDVCGFPVCRPCY
                     EYERKDGSQACPQCKTKYKRHKGSPPILGDESDDVDADDASDVNYPTSGNQDHKHKIA
                     ERMLTWRMNSGRNDDIVHSKYDSGEIGHPKYDSGEIPRIYIPSLTHSQISGEIPGASP
                     DHMMSPVGNIGRRGHPFPYVNHSPNPSREFSGSLGNVAWKERVDGWKMKDKGAIPMAN
                     GTSIAPSEGRGVGDIDASTDYNMEDALLNDETRQPLSRKVPISSSRINPYRMVIVLRL
                     IVLCIFLHYRITNPVRNAYPLWLLSVICEIWFALSWILDQFPKWSPINRETYLDRLAL
                     RYDREGEPSQLAPVDIFVSTVDPMKEPPLVTANTVLSILAVDYPVDKVSCYVSDDGAA
                     MLTFDALAETSEFARKWVPFCKKYSIEPRAPEWYFAQKIDYLKDKVQASFVKDRRAMK
                     REYEEFKVRVNALVAKAQKVPEEGWIMQDGTPWPGNNTRDHPGMIQVFLGHSGGLDTE
```

```
                     GNELPRLVYVSREKRPGFQHHKKAGAMNALVRVSAVLTNGQYLLNLDCDHYINNSKAL
                     REAMCFLMDPNLGRRVCYVQFPQRFDGIDRNDRYANRNTVFFDINLRGLDGLQGPVYV
                     GTGCVFNRTALYGYEPPIKQKRPGYFSSLCGGRKKTKKSKEKSTEKKKSHKHVDSSVP
                     VFNLEDIEEGIEGSGFDDEKSLLMSQMSLEKRFGQSSVFVASTLMEYGGVPQSATPES
                     LLKEAIHVISCGYEDKSDWGTEIGWIYGSVTEDILTGFKMHARGWRSIYCMPKRPAFK
                     GSAPINLSDRLNQVLRWALGSVEILFSRHCPIWYGYGGRLKFLERFAYINTTIYPLTS
                     IPLLLYCILPAICLLTGKFIIPEISNFASIWFISLFLSIFATGILEMRWSGVGIDEWW
                     RNEQFWVIGGISAHLFAVFQGLLKVLAGIDTSFTVTSKASDEEGDFAELYMFKWTTLL
                     IPPTTILIINLVGVVAGISYAINSGYQSWGPLFGKLFFAFWVIVHLYPFLKGLMGRQN
                     RTPTIVVVWAILLASIFSLLWVRIDPFTTRVTGPDTQKCGINC"
     gene            complement(21685. .24525)
                     /gene="OSJNBa0059E14.6"
                     /note="similar to polygalacturonase precursor GB:AC26512
                     GI:3320462 (Cucumis melo)"

  Query Match             87.0%;  Score 17.4;  DB 8;  Length 154555;
  Best Local Similarity   94.7%;  Pred. No. 1.5e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAAC 20
              ||||||||||||| ||||||
Db      26680 TGATCGCCAAACAAGCAAC 26698


RESULT 4
AC135242
LOCUS       AC135242               176534 bp    DNA     linear   HTG 28-FEB-2003
DEFINITION  Oryza sativa (japonica cultivar-group) chromosome 11 clone
            OSJNBa0070D14, *** SEQUENCING IN PROGRESS ***, 6 ordered pieces.
ACCESSION   AC135242
VERSION     AC135242.4  GI:28604237
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 176534)
  AUTHORS   Swain,S.C., Sureshbabu,K., Singh,A., Pal,S., Gaikwad,K.,
            Ghazi,I.A., Dixit,A., Yadav,M., Srivastava,S., Bhargava,A.,
            Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and
            Singh,N.K.
  TITLE     Genomic sequence for Oryza sativa chromosome 11 Clone OSJNBa0070D14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 176534)
  AUTHORS   Swain,S.C., Sureshbabu,K., Singh,A., Pal,S., Gaikwad,K.,
            Ghazi,I.A., Dixit,A., Yadav,M., Srivastava,S., Bhargava,A.,
            Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and
            Singh,N.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-OCT-2002) IIRGS, NRC on Plant Biotechnology, Indian
            Agricultural Research Institute, LBS Centre, New Delhi, Delhi
            110012, India
REFERENCE   3  (bases 1 to 176534)
  AUTHORS   Swain,S.C., Sureshbabu,K., Singh,A., Pal,S., Gaikwad,K.,
            Ghazi,I.A., Dixit,A., Yadav,M., Srivastava,S., Bhargava,A.,
            Pal,A.K., Dalal,V., Batra,K., Sharma,T.R., Mohapatra,T. and
            Singh,N.K.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-FEB-2003) IIRGS, NRC on Plant Biotechnology, Indian
            Agricultural Research Institute, LBS Centre, New Delhi, Delhi
            110012, India
COMMENT     On Feb 28, 2003 this sequence version replaced gi:27502456.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    49567: contig of 49567 bp in length
            *    49568    49667: gap of unknown length
            *    49668    97492: contig of 47825 bp in length
            *    97493    97592: gap of unknown length
            *    97593   145324: contig of 47732 bp in length
            *   145325   145424: gap of unknown length
            *   145425   148804: contig of 3380 bp in length
            *   148805   148904: gap of unknown length
            *   148905   165731: contig of 16827 bp in length
            *   165732   165831: gap of unknown length
            *   165832   176534: contig of 10703 bp in length.
FEATURES    .           Location/Qualifiers
     source             1. .176534
                        /organism="Oryza sativa (japonica cultivar-group)"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:39947"
                        /chromosome="11"
                        /clone="OSJNBa0070D14"
                        /note="(japonica cultivar-group)"
BASE COUNT    50370 a  38483 c  37230 g  49875 t    576 others
ORIGIN

  Query Match             85.0%;  Score 17;  DB 2;  Length 176534;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 ATCGCCAAACGAGCAAC 20
              |||||||||||||||||
Db      80421 ATCGCCAAACGAGCAAC 80437


RESULT 5
AY058358
LOCUS       AY058358               1541 bp    mRNA    linear   INV 16-OCT-2001
DEFINITION  Drosophila melanogaster GH12788 full length cDNA.
ACCESSION   AY058358
VERSION     AY058358.1  GI:16182835
KEYWORDS    FLI_CDNA.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1541)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
            Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
            Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanenavong,S., Wan,K.,
            Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
            This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone.  However, there are
            artifacts associated with the generation of cDNA clones that may
            have not been detected in our initial analyses such as internal
            priming, priming from contaminating genomic DNA, retained introns
            due to reverse transcription of unspliced precursor RNAs, and
            reverse transcriptase errors that result in single base changes.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our Web site
            (http://fruitfly.berkeley.edu) or send email to
            cdna@fruitfly.berkeley.edu.
FEATURES            Location/Qualifiers
     source         1. .1541
```

```
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /map="48E2-48E2"
     gene            1. .1541
                     /gene="CG17509"
                     /note="alignment with genomic scaffold AE003823"
                     /db_xref="FLYBASE:FBgn0033676"
     CDS             28. .1275
                     /gene="CG17509"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="GH12788p"
                     /protein_id="AAL13587.1"
                     /db_xref="GI:16182836"
                     /db_xref="FLYBASE:FBgn0033676"
                     /translation="MIARRIVKELLIQEVPAQRDYELPEDSDWCAQEKLPPDTLCKLD
                     ALKAMARWLLGLRTDEHAAQKTFRMLAAFVNQRGDLLGQNRLCGAEKSWLRLGAACAM
                     LKVCEQKGVGDQYSAEQYLQLSQLMADPVPEVREIFARKLHKGLSRSLPRNCLPLDFM
                     GLYVLAGLETERKLQDLVRHYAETDVNKRREYLKTVAMTSPDSSTESQSLHILPDYML
                     AFAIPVLVHDPRFTNHEDYVQLRKMEKCLRFILEPLMAKRETFVHSFYKQLLQLIKHR
                     EFSLGSDKRDNYKMWALCDLAMYIIDSKFSPFDGNTSTFSMPLALPEMYYKEPAVANF
                     QNNDVYIPLDVYTLGAKSTSKAAATAMTTSRAAVAPKRPAEQSIMDDENPQENNLFDN
                     IRAADTTEPMAKRTRAGAASAKS"
BASE COUNT      435 a    376 c    390 g    340 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 3;  Length 1541;
  Best Local Similarity   90.0%;  Pred. No. 2.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||| |||||||||| |||
Db        787 CTGATGGCCAAACGAGAAAC 806


RESULT 6
F21B23
LOCUS       F21B23                 71284 bp    DNA     linear   PLN 27-JUL-2000
DEFINITION  Arabidopsis thaliana BAC F21B23.
ACCESSION   AF262038
VERSION     AF262038.1  GI:9502148
KEYWORDS    .
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 71284)
  AUTHORS   Washington University Genome Sequencing Center.
  TITLE     The A. thaliana Genome Sequencing Project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 71284)
  AUTHORS   Wilson,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAY-2000) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     Submitted by:
                    Genome Sequencing Center
                    Department of Genetics, Washington University,
                    St. Louis, MO 63108, USA
                    e-mail: rwilson@watson.wustl.edu


            MAPPING: Clones were assigned to the YAC map by hybridization by
            M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted
            by M. Marra, WashU, to pick the best candidates for sequencing.


            NOTICE:  This sequence may not be the entire insert of this clone.
            It may be shorter because we only sequence overlapping sections
            once, or longer because we provide a small overlap between
            neighboring submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality
            >= 30); an attempt was made to resolve all sequencing problems,
            such as compressions and repeats; all regions were covered by
            sequence from more than one m13 subclone.


            NOTES:

            Coding sequences below are predicted from computer analysis, using
            the program Genefinder(P. Green and L. Hillier, ms in preparation).
FEATURES             Location/Qualifiers
     source          1. .71284
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="V"
                     /map="unknown"
                     /clone="F21B23"
     gene            2429. .8909
                     /gene="F21B23.3"
     CDS             join(2429. .2626,2668. .2775,3508. .3920,4187. .4538,
                     4709. .4932,5136. .5331,5530. .5597,5670. .5844,5931. .5978,
                     6028. .6081,6258. .6437,6567. .6679,6803. .6985,7177. .7234,
                     7456. .7515,7615. .7656,7752. .7879,7977. .8112,8243. .8278,
                     8454. .8570,8662. .8724,8808. .8909)
                     /gene="F21B23.3"
                     /note="contains similarity to Drosophila melanogaster
                     BcDNA.GH03694 (GB:AAD55412); coded for by A. thaliana cDNA
                     H36877; coded for by A. thaliana cDNA N96846; coded for by
                     A. thaliana cDNA AA651310"
                     /codon_start=1
                     /protein_id="AAF88009.1"
                     /db_xref="GI:9502156"
                     /translation="MYMAYGWPQVIPLLPGSCPSSQRVVYLKLAGRLLLVVSPSHLEL
                     WGSSQVLYHHLSAFPSNWSLELQRVRLGKYMRDDKSVREEGENLQAVWSPDAKLIAVL
                     FGGAFSIGSHPSDSNDDRLLSYTLGNGLVSGVASPTLASDDKFSTNPAIVQLELCTRS
                     KLLFVLNSDGQLVVCSVNKKGLKYTESIRAEKKVGGDAVCASVASEQQILAVGTRKGM
                     VELYDLSHSISLLRTVSLHDWGYSADYTGPVNNIAWTPDNSAFAVGWKSRGLAVWSVS
                     GCRLMSTVRQIGLTSTSSPKINPKQDCKYEPLMSGTSAIQWDEYGYRLFATEEASYDR
                     ILAFSFGKCCLNRGVSGKTYVRQVSYISQNWPVQHVAASEDGKYLAVAGLHGLILYDI
                     RFKKWRVFGDVSQEQQIHCKGLLWLGKIVVICNYIEASETYELLFYPRYHLDQSSLLC
                     RKVLLGKPMVMDVYQDYILVSYLPFFIHVYHVKIYGELTPSSKADLQHLREGELDNDN
                     LSSDLSDREPSRCLILRGNGELSLLDLVDGRERELTDSVELFWVTCGQSEEKTNLVEE
                     VSWLDYGHRGMQGCPAPKCLTVTLYLIPVWYPSLGDDPFMQEDFLQLDPELEFDREVY
                     PLGLLPNVGVVVGVSQRMSFSASAEFACFEPTPQAQTILHCLLRHLLQRDKNEEALLL
                     AQLSAEKPHFSHCLEWLLFTVFDAEISRPNPNRNQISGPGHLKKLSLLRKACDLIKKF
                     PEYYDVVVNVARKTDARHWADLFSAAGISTTLFEDCFQRRWYRTAACYILVIAKLEGV
                     AVSQYCALRLLQATLDESLYDLAGELVRFLLRSGRDIEQAPTESDSLSPKLLGFLIFG
                     SSHKKSSLDKSSSFKEQSPHVASVKSILESHASYLMSGKELSKLVAFVKGTQFDIVNF
                     AAGLELIGQKLQMSELQNRLDAEFLLAQMCSVKFKEWIVVLATLLQRSEVLYDIFRYD
                     LRLWKAYSVTLESHLAFAHYHDLLQILEAKLSATSREESNRGSIS"
     gene            complement(10977. .11830)
                     /gene="F21B23.5"
     CDS             complement(join(10977. .11226,11408. .11588,11647. .11830))
                     /gene="F21B23.5"
                     /note="contains similarity to aminocyclopropanecarboxylate
                     synthases"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF88004.1"
                     /db_xref="GI:9502151"
                     /translation="MSQGACENQLLSKLALSDKHGEASPYFHGWKAYDNNPFHPTHNP
                     QGVIQMGLAENQVRLLIIYFIVICFQLCSDLIKEWIKENPHASICTAEGIDSFSDIAV
                     FQDYHGLKQFRQVFISFLIRTIICFQIMFCLGTENKQAIATFMERARGGRVRFEAERV
                     VMSGGATGANETIMFCLADPGDAFLVPTPYYAAYVYIHTLLVFK"
     gene            complement(14503. .16686)
                     /gene="F21B23.6"
     CDS             complement(14503. .16686)
```

```
                     /gene="F21B23.6"
                     /note="similar to a large family of Arabidopsis thaliana
                     salt inducible protein-like proteins; contains similarity
                     to Pfam family PF01535 (Domain of unknown function),
                     score=340.5, E=1.9e-98, N=2"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF88005.1"
                     /db_xref="GI:9502152"
                     /translation="MSIMLSISRRQNSYILLNHSRFLRRFSYDVDPRPEIKSESQEFV
                     VVKFVKTLQNTPQHDWASSESLSALVVSSSSASPLVFSQITRRLGSYSLAISFFEYLD
                     AKSQSLKRREESLSLALQSVIEFAGSEPDPRDKLLRLYEIAKEKNIPLTIVATKLLIR
                     WFGRMGMVNQSVLVYERLDSNMKNSQVRNVVVDVLLRNGLVDDAFKVLDEMLQKESVF
                     PPNRITADIVLHEVWKERLLTEEKIIALISRFSSHGVSPNSVWLTRFISSLCKNARAN
                     TAWDILSDLMKNKTPLEAPPFNALLSCLGRNMDISRMNDLVLKMDEVKIRPDVVTLGI
                     LINTLCKSRRVDEALEVFEQMRGKRTDDGNVIKADSIHFNTLIDGLCKVGRLKEAEEL
                     LVRMKLEERCVPNAVTYNCLIDGYCRAGKLETAKEVVSRMKEDEIKPNVVTVNTIVGG
                     MCRHHGLNMAVVFFMDMEKEGVKGNVVTYMTLIHACCSVSNVEKAMYWYEKMLEAGCS
                     PDAKIYYALISGLCQVRRDHDAIRVVEKLKEGGFSLDLLAYNMLIGLFCDKNNAEKVY
                     EMLTDMEKEGKKPDSITYNTLISFFGKHKDFESVERMMEQMREDGLDPTVTTYGAVID
                     AYCSVGELDEALKLFKDMGLHSKVNPNTVIYNILINAFSKLGNFGQALSLKEEMKMKM
                     VRPNVETYNALFKCLNEKTQGETLLKLMDEMVEHLVNQIRSQWRF"
gene                 complement(24152. .26942)
                     /gene="F21B23.7"
CDS                  complement(join(24152. .24287,24374. .24440,24665. .24872,
                     25152. .25208,25625. .25735,26021. .26062,26667. .26942))
                     /gene="F21B23.7"
                     /note="contains similarity to Pfam family PF00076 (RNA
                     recognition motif. (a.k.a. RRM, RBD, or RNP domain)),
                     score=42.1, E=1.3e-08, N=1"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF88006.1"
                     /db_xref="GI:9502153"
                     /translation="MVDSGFPPEHEIAASDKEKVSNLVPLVPTHGIIQSSEAEYCGICS
                     NPIFTVYLDSVGNVAKVEVDEDEVLSRRSGLDDLELVSSRLQKVLEDQMFKFALLIGL
                     GRQKIMREKEPGDGKGYPFVTFRCKDLAAEAIDNLNNTVFREPQNMGRNRGFAFIEYY
                     NHVKALYIKNLPRDITQERLKALFEHHGKILKVVIPPAKPGKEDSRYGFVHYAERTSV
                     MRALKNTERYEIDGHAAGGMSMMPIMLPDGRIRYVLQQPGLAAMPQPPPRPSPPYRGG
                     SGSSSSKQSSDNGRGRSRYNPY"
gene                 complement(27206. .29561)
                     /gene="F21B23.8"
CDS                  complement(join(27206. .27759,28027. .28122,28481. .29561))
                     /gene="F21B23.8"
                     /note="contains similarity to Drosophila
                     melanogasterspectrin beta chain (GB:AAA28399.1) and
                     heparan sulfate-N-deacetylase/N-sulfotransferases"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF88007.1"
                     /db_xref="GI:9502154"
                     /translation="MENGEYVTHKDTSLLKFVSSSSSSSSEEEFRFVSSIQNAILRLD
                     LIPKLPAIGRALLFGYIGLWLLKTVLVYRKSNEVECTELEKEMMRRKMKAWEERDMSE
                     KGTVEVLHKEGLEKPLMSFEKPKFDRNELMSSISKVKGSEKKLELVNSSHVELDFDDK
                     IHEIKVMARRAREIEAGIELNEKEKRDVNKETGDSDEDISIQSQKSLPHDGLTHSVGD
                     DDKDERLGTSTDSENTELSAFAVPMLNGAMVDSGFPNHEMAASDKKKVSNVVPLVPTD
                     GVIQASDVTKDQLSMMKNSTGRKSRVIRSVKEAKEFLSRRSGEKELTQELSQMAQDSD
                     EIFPKQSDEERGVARKHKLVDKNKILELYRSEYNDELEWMKDEKLRDIVFCVRDNELA
                     DGVSVYDPLEKIIPRWKGPSLDKNPEFLNNYHEQREALFSGKAASVSPVKYEEQSSHQ
                     ELSESASSENTLTPSSEITSSQPKIVVEGSDGSVRPGKKSGKEYWQHTKKWSRGFLEL
                     YNAETDPEVKAVMRDMGKDLDRWITEDEIKDAADIMEKLPERNKKFMEKKLNKLKREM
                     ELFGPQAVVS"
gene                 39785. .41443
                     /gene="F21B23.2"
CDS                  join(39785. .39801,40280. .40444,40500. .40569,40634. .40735,
                     40816. .40861,40959. .41072,41156. .41317,41364. .41443)
                     /gene="F21B23.2"
                     /note="contains similarity to Pfam family PF00350 (Dynamin
                     family), score=29.7, E=4e-06, N=2"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF88003.1"
                     /db_xref="GI:9502150"
                     /translation="MSSIKRFLHGSYELSTAEAFLCIPSSLSFVSFSDPSSLSSLPFI
                     TIAVSPLPFSLSLLRSPMRISVMKPVRHQQVLLPGLLKRAGVNSRYWRESKKSLFVAE
                     YNYKRMMLKKKKKKKKLNCIAANQDCEVDPKGDMTKIDLKDQGTNAVETGRSSKVVST
                     VPLHRSIFSPNGCDIQSSFKFMCSTYHLSHLSSCIHTLTNFFLLKMFVIVNLTLVDLP
                     GLTKVAVGNKAMMSSKAEARDQMPLVIQVLQFH"
gene                 complement(47870. .48406)
                     /gene="F21B23.9"
CDS                  complement(join(47870. .48214,48338. .48406))
                     /gene="F21B23.9"
                     /note="contains similarity to human ezrin (GB:AAA61278.1)"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="AAF88008.1"
                     /db_xref="GI:9502155"
                     /translation="MAYVNKENAIKMAEQNLKANSEIHKVSLLDEVNDRFMYLSQARA
                     NAQLIEVLEGGGVLEREKEQVDEWLKDFADAEVNLNRFMAELKDDLKAPAPEPAPLSP
                     GGHRSVESLADEAGITDQAGSLLPAKDNRPSEDLD"
gene                 57055. .59355
                     /gene="F21B23.1"
CDS                  57055. .59355
```

```
  Query Match             84.0%;  Score 16.8;  DB 8;  Length 71284;
  Best Local Similarity   90.0%;  Pred. No. 3.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              |||||||||||  |||||||
Db      18720 CTGATCGCCAATGGAGCAAC 18739
```

```
RESULT 7
AC006837/c
LOCUS       AC006837               91268 bp    DNA     linear   PLN 27-FEB-2002
DEFINITION  Arabidopsis thaliana chromosome 2 BAC F23H14 genomic sequence,
            complete sequence.
ACCESSION   AC006837
VERSION     AC006837.16  GI:20197903
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 91268)
  AUTHORS   Town,C.D., Haas,B.J., Wu,D., Maiti,R., Hannick,L.I., Chan,A.P.,
            Tallon,L.J., Rooney,T., Utterback,T.R., VanAken,S.E.,
            Feldblyum,T.V., White,O. and Fraser,C.M.
  TITLE     Arabidopsis thaliana chromosome 2 BAC F23H14 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 91268)
  AUTHORS   Lin,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
REFERENCE   3  (bases 1 to 91268)
  AUTHORS   Town,C.D. and Kaul,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
COMMENT     On Apr 18, 2002 this sequence version replaced gi:6598619.
            Address all correspondence to:at@tigr.org

            BAC clone F23H14 is from Arabidopsis thaliana chromosome 2
            The orientation of the sequence is from SP6 to T7 end of the BAC
            clone.
            Genes were identified by a combination of several methods: Gene
            prediction programs including Genscan+ (Chris Burge,
            http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHMM (Mark Borodovsky,
            http://genemark.biology.gatech.edu/GeneMark/), GlimmerA (a variant
            of GlimmerM, see Mihaela Pertea,
            http://www.tigr.org/softlab/glimmerm_htm/glimmerm.html, and
            GeneSplicer (Mihaela Pertea and Steven Salzberg, contact
            mpertea@tigr.org), searches of the complete sequence against a
```

peptide database and the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by repeatmasker (Arian Smit, http://ftp.genome.washington.edu/RM/RepeatMasker.html).

FEATURES Location/Qualifiers

source 1. .91268
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="2"
/map="mi398"
/clone="F23H14"

misc_feature 1. .9372
/note="rDNA 10.5 kbp repeat unit"

repeat_region 1. .88
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 1. .85
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

repeat_region 80. .2705
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

repeat_region 85. .2705
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 412. .501
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

repeat_region 414. .503
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 414. .501
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 447. .500
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 449. .502
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

repeat_region 1459. .1829
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 1459. .1519
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

repeat_region 1461. .1829
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

repeat_region 1461. .1519
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 1769. .2139
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

repeat_region 1771. .2139
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 2079. .2139
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

repeat_region 2081. .2139
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

repeat_region 2706. .9364
/rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA repeat"

gene 2706. .4513
/gene="At2g01015"

rRNA 2706. .4513
/gene="At2g01015"
/product="18SrRNA"

repeat_region 2706. .4513
/rpt_family="Rf:18S rDNA contained with the 10.5 kbp rDNA repeat#Rf:ATR0027|X16077 18 S rDNA"

repeat_region 4514. .4781
/rpt_family="Rf:spacer between 18S and 5.8S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0028|X52320|X52322|U43224 intergenic spacer between 18S and 5.8S rDNAs"

gene 4782. .4945
/gene="At2g01025"

rRNA 4782. .4945
/gene="At2g01025"
/product="5SrRNA"

repeat_region 4782. .4945
/rpt_family="Rf:5.8S rDNA contained with the 10.5 kbp rDNA repeat#Rf:ATR0029|X52320|U43225 5.8S rDNA"

repeat_region 4946. .5133
/rpt_family="Rf:Spacer between 5.8S and 25S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0030|X52320|X52322|U43225 intergenic spacer between 5.8S and 25S rDNAs"

repeat_region 5134. .8520
/rpt_family="Rf:25S rDNA contained within the 10.5 kbp rDNA repeat#Rf:ATR0031|X52320|X52322 25S rDNA"

repeat_region 8521. .9364
/rpt_family="Rf:Spacer between 25S and 18S rDNAs part of the 10.5 kbp rDNA repeat#Rf:ATR0032|X15550 intergenic spacer between 25S and 18S rDNAs"

```
     repeat_region   8750. .8829
                     /rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic
                     spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat
                     unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and
                     IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA
                     repeat"
     repeat_region   9195. .9372
                     /rpt_family="Rf:rDNA repeat of 18S rDNA, intergenic
                     spacer, 5.8S rDNA, spacer, 25S rDNA, spacer entire repeat
                     unit is 10.5 kbp, 2 loci of ~3.5 Mbp on chromosomes II and
                     IV, 670 copies#Rf:ATR0026|X52322|X15550 10.5 kbp rDNA
                     repeat"
     repeat_region   9197. .9374

  Query Match             84.0%;  Score 16.8;  DB 8;  Length 91268;
  Best Local Similarity   90.0%;  Pred. No. 3.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||||||||| |||||||
Db      53759 CTGATCGCCAATGGAGCAAC 53740


RESULT 8
NCB14A6/c
LOCUS       NCB14A6                91436 bp    DNA     linear   PLN 18-JAN-2002
DEFINITION  Neurospora crassa DNA linkage group V BAC contig B14A6.
ACCESSION   AL670007
VERSION     AL670007.1  GI:18376186
KEYWORDS    .
SOURCE      Neurospora crassa
  ORGANISM  Neurospora crassa
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
REFERENCE   1
  AUTHORS   Schulte,U., Aign,V., Hoheisel,J., Brandt,P., Fartmann,B.,
            Holland,R., Nyakatura,G., Mewes,H.W. and Mannhaupt,G.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 91436)
  AUTHORS   German Neurospora genome,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JAN-2002) MIPS, Institut fuer Bioinformatik,
            GSF-Forschungszentrum fuer Umwelt und Gesundheit, GmbH,
            Ingolstaedter Landstrasse 1, D-85764 Neuherberg, FRG, E-mail:
            G.Mannhaupt@gsf.de Project Coordinator: Ulrich Schulte, Institute
            of Biochemistry, Heinrich-Heine-University, D-40225 Duesseldorf,
            E-mail: ulrich-schulte@uni-duesseldorf.de
COMMENT     this contig is an assembly of BAC 14A6 from 1 to: 11263, and BAC
            12H14 from 11264 to: 91436rev (strain OR74A); BAC clones are
            available at the Fungal Genetic Stock Center http://www.fgsc.net
            Sequencing was performed by MWG Biotech AG, Ebersberg, Germany,
            http://www.mwgdna.com
            Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of linkage groups II
            and V can be viewed at: http://mips.gsf.de/proj/neurospora.
FEATURES             Location/Qualifiers
     source          1. .91436
                     /organism="Neurospora crassa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:5141"
                     /chromosome="2"
     misc_feature    1. .7884
                     /note="overlap to BAC B13O8 contig, please refer to this
                     entry for analysis and annotation"
     tRNA            5365. .5451
                     /product="tRNA-Gln"
                     /note="tRNA predict as a tRNA- Gln : anticodon ttg"
     gene            5919. .7803
                     /gene="B14A6.010"
     gene            complement(join(5919. .6185,6494. .7692,7761. .7803))
                     /gene="B14A6.010"
     CDS             complement(join(5919. .6185,6494. .7692,7761. .7803))
                     /gene="B14A6.010"
                     /note="match of 76 ESTs
                     Contains ATP/GTP-binding site motif A (P-loop)
                     [GGGGVGKT] [ATGGLGKT]
                     contains EST gb:AW717407, AW708782"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAD21303.1"
                     /db_xref="GI:18376187"
                     /translation="MPTATTVKQPSAAVSSTPKVKKTTTKTETTDGPKPVKEVKHRSS
                     STKPKSTPLAETTDNVVKKKPKPAPLDENKIATTDHAVKKKPTTSTSTPEKKKARVHS
                     SSTYPREVVSSSSTKKPSTSGHTHAHTSTTKTSSHRPHTHLEESSKHESTSSHRPGRR
                     HHAHHHHNPDGSLTRHGARATSSSIAKPLAPAKAFSPALDAEGNPKPNLRPATMTAHI
                     KRAVRKSEAMNPSKPTKVEPETALPKVHVPTPQTARQKATAAADVEATKRAQKVAAAA
                     QEEQAQWDQQQQQQQQETQKPKKKEKKYAPEAMIPTSRFVGNVGKTAGAALDTTGQVA
                     KGVTDTVGRTAEGALGGGGVGKTVGAAAGGVGDTLGAATGGLGKTVEGATGGLAKGGP
                     LGAVGGLGEGLGQTVGGVGKGLGDTVGGTVGGLGDTVGGPLGGVVGGLGRGVGGAVTG
                     ITGGLGDGVGKIGQGDLLGGVGDVVGGVTNGVGGLLGGVLGGLGGGGGGQRR"
     exon            complement(5919. .6185)
                     /gene="B14A6.010"
                     /number=1
     intron          complement(6186. .6493)
                     /gene="B14A6.010"
                     /number=1
     exon            complement(6494. .7692)
                     /gene="B14A6.010"
                     /number=2
     intron          complement(7693. .7760)
                     /gene="B14A6.010"
                     /number=2
     exon            complement(7761. .7803)
                     /gene="B14A6.010"
                     /number=3
     gene            10983. .12122
                     /gene="B14A6.020"
     CDS             10983. .12122
                     /gene="B14A6.020"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAD21304.1"
                     /db_xref="GI:18376188"
                     /translation="MSNQRNPSHSLLDVNSFCSDDQKSVQEARNHSTSSNAPTNHAYV
                     IPAQRRAMTPSDTPSVSAALAILDWRRKAQTPTSNTRSATEPVLVTRSPSNSFRPMDS
                     VQGLEPTFTAEEKVRAGFSWRYKGDIRNPKNSSLGIAPDDNCALFLTNIPAGTTPKIV
                     LDALGIHAPFGRVYATSIAPPTDGRYKSACAKITMFCREEAVHVFQFITGGNLKIAGR
                     TIAVVWNSNRSAPQTDLPHSASRVLEILGPIEVVDIHRLSRFLGANIKFDTQDVTTVR
                     EDKDERLLRWTFCSFRAQAQVARKALRAEWPGLHIRWGIDPMAVSVNMVPTAMSTMLP
                     LFPQQRAPVATSEVSHTPAGDLISFDSDVEEGLDSKEEKSDIYDF"
     exon            10983. .12122
                     /gene="B14A6.020"
                     /number=1
     tRNA            13287. .13358
                     /product="tRNA-Thr"
                     /note="tRNA predict as a tRNA- Thr : anticodon agt"
     gene            13804. .15397
                     /gene="B14A6.030"
     gene            complement(join(13804. .15031,15093. .15397))
                     /gene="B14A6.030"
     CDS             complement(join(13804. .15031,15093. .15397))
                     /gene="B14A6.030"
                     /note="strong similarity to benzoate 4-monooxygenase (EC
                     1.14.13.12), Aspergillus niger, PIR:S12015
                     Contains Cytochrome P450 cysteine heme-iron ligand
                     signature [FSHGPRSCVG]"
                     /codon_start=1
                     /product="probable benzoate 4-monooxygenase cytochrome
                     P450"
                     /protein_id="CAD21305.1"
                     /db_xref="GI:18376189"
                     /translation="MAIISLLMSPWAPVVLLAGVAFYYLVPYFVTYSALRKIPSPFPA
                     QFTDLWLLSVCRRGNRYQRVDELHKKLGPVVRIQPNHVSICDDAAIPTIYGHGNGFLN
                     DFYDAFVSIRRGLFNTRDRAEHTRKRKIVSHTFSAKSVQQFEPYMHSNLELFVKQWDS
                     MIKNSKNPDKAAHLDCLEWFNYLAFDVIGDLSFGQPFGMLSSGADMAEIRSSPDAAPI
                     YAPAIEILNRRGEVSATLGIHPALKPFAKYLPDPFFTKGLAAVENLAGIAIACVKSRL
```

```
                DNPPPVTRKDLLQRLMEGRDEKGEPLGREELTAEALTQLIAGSDTTSNSSCALLFHAV
                RTPGVMQKLQAELDANIPPEVDVPTYDMVKELPYLEAVINEVLRFHSTSGIGLPRQIP
                HDASQGVHIQGYYLPPGTVLSVPTYSIHHSKEIWGPDADEFKPERWERLTARQKNAFI
                PFSHGPRSCVGRNVAEMEMKLIVATWARRYEVKLLQDYMDTREGFLRKPLGLKVGLKL
                RK"
exon            complement(13804. .15031)
                /gene="B14A6.030"
                /number=1
intron          complement(15032. .15092)
                /gene="B14A6.030"
                /number=1
exon            complement(15093. .15397)
                /gene="B14A6.030"
                /number=2
gene            17314. .19917
                /gene="B14A6.040"
gene            complement(17314. .19917)
                /gene="B14A6.040"
CDS             complement(17314. .19917)
                /gene="B14A6.040"
                /note="similarity to glucan 1, 4-alpha-glucosidase (EC
                3.2.1.3), YIR019c, Saccharomyces cerevisiae, PIR:S48478"
                /codon_start=1
                /product="related to glucan 1, 4-alpha-glucosidase"
                /protein_id="CAD21306.1"
                /db_xref="GI:18376190"
                /translation="MDDPWGSPWATADASTDLGHSLTRRRSDLEPPKRALLSASNSPR
                LPAITAPSPWADAEDGFGGWAAPDGLTTQTATTAQSAWMGGWRSRSPSLAAPSRDNGF
                RSRSPSLVAPSRENSWRSRSPSLIAPSRDNGFRSRSPSIAASVRDEEFAKSNPIAWPD
                TIASSNSPIAPVLRQPSPDPWAAEHSHTPHNDDNITPRLVINLPQSIAEEEDDVDKGK
                LSELNPDWADEEKDASKDGACEEEQSEPTWAKAETEAEVVQRVPLRISVGSTVNSLPT
                RSTSPSGDDTDHGDERQDSPITSIDDDSHARPRLARKSSSIVQKPVEEFDKTAQVMSQ
                GPPVIRRAESKSSLGPHHHDVSEEASDFGDFEDAQVKEPPSSQTPEPPSTPSEVQKPS
                SQHGSSPTTSPTQSMSSQLGHMNELAAKFGALNFDIDLSNVDKMFETSTNMKPLFPYA
                DVSDHIITDSFNSISERKAWYRISRFGSSRKHNAGDDENYRRVAWPTSTVCQETLQTV
                RRWMEEDSIAGRVSLGGGISKTQKNMFGWDSSAEPVTLDAIFGKRAVPTRPSSVQPLQ
                PPSAFNLVDHSPLGSTAVNSLKMQPHRPASLQLPPAATFSWSSDFSASAGKPPLTASL
                PLTPSQPAQGPFKVTPMDLGMSSISATAAAEDNDDDWGEMVSSPSVTKATVGGFPDFN
                SPVSTLTSTRGPNPGTTTAEYMSQSDHVPVSLTLTSNLENNTTVPVTSSAADPWASAD
                FSFFELPSQQQIGPTKSSVTPSRPSSMIATSFAPTAPGSSTHSFDLPRVTTPISSASS
                VIAPHSSMTRNSSPIAAKGTITSAISSPIARQVTPTQATGDLQQGEPTEDEVTVSSVL
                GGLPDLSYMLR"
exon            complement(17314. .19917)
                /gene="B14A6.040"
                /number=1
gene            20612. .21612
                /gene="B14A6.050"
CDS             join(20612. .20644,20731. .21612)
                /gene="B14A6.050"
                /note="similarity to putative nucleotide binding protein,
                Homo sapiens, TREMBL:AF118394_1
                Contains ATP/GTP-binding site motif A (P-loop) [GKGGVGKS]"
                /codon_start=1
                /product="related to putative nucleotide binding protein
                (NUPB)"
                /protein_id="CAD21307.1"
                /db_xref="GI:18376191"
                /translation="MSLAKVKHIVLVLSGKGGVGKSSVTTQLALSLSLAGHSVGVLDV
                DLTGPSIPRMFGIEDAKVTQAPGGWLPITVHEADPSAGVGSLRVMSLGFLLPKRGDAV
                VWRGPKKTAMVRQFLSDVFWDETDYLLIDTPPGTSDEHISLAENLLQKARPGQLAGAV
                VVTTPQAVATADVRKELNFCTKTNIRVLGVVENMCGFVCPNCSECTNIFMSGGGEVMA
                NDFGVRFLGRVPIDPQFLVLIETGKRPTYPAGTTVDGKDISTPAGASTSEEEEVKDGS
                RLVHKYKDCSLAPIFSKITADVISAVQQ"
exon            20612. .20644
                /gene="B14A6.050"
                /number=1
intron          20645. .20730
                /gene="B14A6.050"
                /number=1
exon            20731. .21612
                /gene="B14A6.050"
                /number=2
gene            23875. .25386
                /gene="B14A6.060"
CDS             23875. .25386
                /gene="B14A6.060"
                /note="identical to TREMBL:AF151380_1
                Contains Fumarate lyases signature [GSAYMVGKMN]"
                /codon_start=1

  Query Match             84.0%;  Score 16.8;  DB 8;  Length 91436;
  Best Local Similarity   90.0%;  Pred. No. 3.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||| |||||||| ||||
Db      44965 CTGATGGCCAAACGAACAAC 44946


RESULT 9
ATAC010927/c
LOCUS       ATAC010927            96232 bp    DNA     linear   PLN 30-OCT-2002
DEFINITION  Arabidopsis thaliana chromosome III BAC T22K18 genomic sequence,
            complete sequence.
ACCESSION   AC010927
VERSION     AC010927.5  GI:12408726
KEYWORDS    HTG.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 96232)
  AUTHORS   Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.,
            Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R., Barnstead,M.E.,
            Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
  TITLE     Arabidopsis thaliana chromosome III BAC T22K18 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 96232)
  AUTHORS   Lin,X. and Kaul,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-SEP-1999) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org
REFERENCE   3  (bases 1 to 96232)
  AUTHORS   Lin,X.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
            Medical Center Dr., Rockville, MD 20850, USA
COMMENT     On Jan 24, 2001 this sequence version replaced gi:12280869.
            Address all correspondence to:
            Xiaoying Lin
            The Institute for Genomic Research
            9712 Medical Center Dr.
            Rockville, MD 20850, USA
            e-mail: xlin@tigr.org
            BAC clone T22K18 is from Arabidopsis chromosome III and is near the
            molecular marker CIC5D8.
            The orientation of the sequence is from SP6 to T7 end of the BAC
            clone.
            Genes were identified by a combination of three methods: Gene
            prediction programs including GRAIL (available by anonymous ftp
            from arthur.epm.ornl.gov), Genefinder (Phil Green, University of
            Washington), Genscan (Chris Burge,
            http://gnomic.stanford.edu/~chris/GENSCANW.html), and NetPlantGene
            (http://www.cbs.dtu.dk/netpgene/cbsnetpgene.html), searches of the
            complete sequence against a peptide database and the Arabidopsis
            EST database at TIGR (http://www.tigr.org/tdb/at/at.html).
            Annotated genes are named to indicate the level of evidence for
            their annotation. Genes with similarity to other proteins are named
            after the database hits. Genes without significant peptide
            similarity but with EST similarity are named as 'unknown' proteins.
            Genes without protein or EST similarity, that are predicted by more
            than two gene prediction programs over most of their length are
            annotated as 'hypothetical' proteins. Genes encoding tRNAs are
            predicted by tRNAscan-SE (Sean Eddy,
            http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are
```

```
        identified by repeatmasker (Arian Smit,
        http://ftp.genome.washington.edu/RM/RepeatMasker.html). Regions of
        genomic sequence that are not annotated as genes but have predicted
        exons by GRAIL are annotated as misc features.
FEATURES             Location/Qualifiers
     source          1. .96232
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="III"
                     /map="CIC5D8"
                     /clone="T22K18"
     misc_feature    1. .708
                     /note="overlap with BAC clone F14P13
                     (AC009400:87230. .87937)."
     misc_feature    complement(48. .91)
                     /note="exon predicted by xgrail, quality marginal"
     misc_feature    complement(437. .495)
                     /note="exon predicted by xgrail, quality good_shadowexon"
     gene            <877. .>3949
                     /gene="T22K18.1"
                     /note="similar to folylpolyglutamate synthetase
                     GB:AAA87568 [Homo sapiens]"
     mRNA            join(<877. .929,1009. .1069,1403. .1486,1588. .1644,
                     1737. .1797,1880. .1926,2058. .2148,2480. .2619,2713. .2955,
                     3152. .3230,3312. .3467,3548. .3744,3890. .>3949)
                     /gene="T22K18.1"
     CDS             join(877. .929,1009. .1069,1403. .1486,1588. .1644,1737. .1797,
                     1880. .1926,2058. .2148,2480. .2619,2713. .2955,3152. .3230,
                     3312. .3467,3548. .3744,3890. .3949)
                     /gene="T22K18.1"
                     /codon_start=1
                     /product="putative folylpolyglutamate synthetase"
                     /protein_id="AAF04408.1"
                     /db_xref="GI:6143861"
                     /translation="MFTSPHLIDVRERFRIDGLDISEEKFLQYFWECWKLLKEKAVDG
                     LTMPPLFQFLTVLAFKIFVCEKVDVAVIEVGLGGKLDSTNVIQKPVVCGIASLGMDHM
                     DILGNTLADIAFHKAGIFKPQIPAFTVPQLSEAMDVLQKTANNLEVTIVAPLEPKKLD
                     GVTLGLSGDHQLVNAGLAVSLSRCWLQRTGNWKKIFPNESKETEIPVAFCRGLATARL
                     HGRAQVVHDVVSDPQDSSDSMETPCGDLIFYLDGAHSPESMEACGRWFSSAVRGDKSL
                     STAILLFNCMEVRDPQVLLPKLVTTCASSGTHFSRALFVPSMSTYNKVISGASAIPSD
                     TRRKDLTWQFRLQRLWEKSIQGTDAGLDHTLKPDGITALPPHDFLCGDAPQCGGPAGT
                     PVTSSAVMPSLPLTINWLRDCVRRNPSLKLEVLVTGSLHLVGDVLRLLKR"
     misc_feature    4078. .4106
                     /note="exon predicted by xgrail, quality marginal"
     gene            <4481. .>6035
                     /gene="T22K18.2"
                     /note="predicted by genscan+"
     mRNA            join(<4481. .4703,4985. .5477,5564. .5732,5817. .>6035)
                     /gene="T22K18.2"
     CDS             join(4481. .4703,4985. .5477,5564. .5732,5817. .6035)
                     /gene="T22K18.2"
                     /note="hypothetical protein"
                     /codon_start=1
                     /protein_id="AAF04409.1"
                     /db_xref="GI:6143862"
                     /translation="MKKPSLFQIIIIVLSIPTTTGRTVGNLRVREGSPFKIAIFADLH
                     FGEDTWTDWGPGQDVNSVNVMSAVLDAETPDFVVYLGDVVTANNIAIQNASLFWDKAI
                     SPTRDRGIPWATLFGNHDDASFVWPLDWLSSSGIPPLRCPAASDDDGCTFRGTTRVEL
                     IQEEIKSSNALSYSMISPKELWPSVSNYVLLVESSDHSKPPVALLYFLDSGGGSYPEV
                     ISNAQVEWFKTKSNTLNPYLRIPELIFWHIPSKAYKKVAPRLWITKPCVGSINKEKVV
                     AQEAENGMMRVLENRSSVKAVFVGHNHGLDWCCPYKDKLWLCFARHTGYGGYGNWPRG
                     SRILEISEMPFRIKTWIRMEDGSVHSEVNLTYD"
     misc_feature    complement(6053. .6230)
                     /note="exon predicted by xgrail, quality good_shadowexon"
     gene            complement(<6468. .>8553)
                     /gene="T22K18.3"
                     /note="similar to RecA protein GB:BAA78779
                     [Rhodopseudomonas palustris]"
     mRNA            complement(join(<6468. .6744,6877. .7016,7118. .7264,
                     7342. .7514,7741. .7828,7938. .8155,8232. .8295,8491. .>8553))
                     /gene="T22K18.3"
     CDS             complement(join(6468. .6744,6877. .7016,7118. .7264,
                     7342. .7514,7741. .7828,7938. .8155,8232. .8295,8491. .8553))
                     /gene="T22K18.3"
                     /codon_start=1
                     /product="putative RecA protein"
                     /protein_id="AAF04410.1"
                     /db_xref="GI:6143863"
                     /translation="MGRLSWASPIQRFRFFSYLSQLNGRRSVLACSGYENRYLSSLVE
                     ASDCELDEVPDDRKVAEKDTALHLALSQLSGDFDKDSKLSLQRFYRKRRVSVISTGSL
                     NLDLALGVGGLPKGRMVEVYGKEASGKTTLALHIIKEAQKLGGYCAYLDAENAMDPSL
                     AESIGVNTEELLISRPSSAEKMLNIVDVLTKSGSVDVIVVDSVAALAPQCELDAPVGE
                     RYRDTQSRIMTQALRKIHYSVGYSQTLIVFLNQVRSHVKSNMHFPHAEEVTCGGNALP
                     FHAAIRLKMIRTGLIKTANKVYGLNVCVQVVKNKLAPGKKKSELGIHFGHGFYVEREV
                     LELACEHGVILREGTSYFIEGEVIEGKDAAEKYLVENKEALDTVVAILRNQLFKM"
     misc_feature    complement(8596. .8676)
                     /note="exon predicted by xgrail, quality good_shadowexon"
     repeat_region   complement(8803. .8900)
                     /rpt_family="(TAAAA)n"
     repeat_region   complement(8805. .8840)
                     /rpt_family="POLY_A"
     misc_feature    8974. .9235
                     /note="exon predicted by xgrail, quality excellent"
     misc_feature    9303. .9423
                     /note="exon predicted by xgrail, quality excellent"
     misc_feature    complement(9461. .9610)
                     /note="exon predicted by xgrail, quality
                     marginal_shadowexon"
     gene            <10379. .>12415
                     /gene="T22K18.4"
                     /note="predicted by genscan+, est match"
     mRNA            join(<10379. .10619,10701. .10790,10871. .10941,11030. .11131,
                     11228. .11332,11695. .11830,12151. .12260,12341. .>12415)
                     /gene="T22K18.4"
     CDS             join(10379. .10619,10701. .10790,10871. .10941,11030. .11131,
                     11228. .11332,11695. .11830,12151. .12260,12341. .12415)
                     /gene="T22K18.4"
                     /note="unknown protein"
                     /codon_start=1
                     /protein_id="AAF04411.1"
                     /db_xref="GI:6143864"
                     /translation="MMMISSSSITSSLSLLSSSPEKLPLINPIQRCPITYSGFRTASV
                     NRAIRRQPQSPAVSATESRVSLVLALASQASSVSQRLLADLAMETAKYVFPKRFDSST
                     NLEEAFMSVPDLETMNFRVLFRTDKYEIRQVEPYFVAETIMPGETGFDSYGASKSFNV
                     LAEYLFGKNTIKEKMEMTTPVVTRKVQSVGEKMEMTTPVITSKAKDQNQWRMSFVMPS
                     KYGSNLPLPKDPSVKIQQVPRKIVAVVAFSGYVTDEEIERRERELRRALQNDKKFRVR
                     DGVSFEVAQYNPPFTLPFMRRNEVSLEVENKED"
     misc_feature    complement(12526. .12557)
                     /note="exon predicted by xgrail, quality
                     marginal_shadowexon"
     gene            complement(12647. .13454)
                     /gene="T22K18.5"
                     /note="predicted by genscan+, multiple est matches"
     mRNA            complement(12647. .13454)
                     /gene="T22K18.5"
     CDS             complement(12861. .13364)
                     /gene="T22K18.5"
                     /note="unknown protein"
                     /codon_start=1
                     /protein_id="AAF04412.1"
                     /db_xref="GI:6143865"
                     /translation="MEKKVIKIMRNDGKVVEYRGPMKVHHILTQFSPHYSLFDSLTNN
                     CHLHPQAKLLCGRLYYLLPQETNSIKHMKKTMKKVRFANPEVEKEEQEEDRLTDCCDN
                     TKEKTNGVVRVKMVVSKQELEKLLQGGSVHEMVYRTLAKQHLCDDDDECHKEGWRPLL
                     DSIPETD"
     repeat_region   complement(13084. .13157)
                     /rpt_family="(GAA)n"
     repeat_region   complement(13387. .13465)
                     /rpt_family="(CAAA)n"
     misc_feature    13913. .13972
                     /note="exon predicted by xgrail, quality good_shadowexon"
     misc_feature    complement(14089. .14181)
                     /note="exon predicted by xgrail, quality marginal"
```

```
  Query Match             84.0%;  Score 16.8;  DB 8;  Length 96232;
  Best Local Similarity   90.0%;  Pred. No. 3.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              |||||||||||  |||||||
Db      19925 CTGATCGCCAATGGAGCAAC 19906


RESULT 10
AC007475
LOCUS       AC007475              173613 bp    DNA     linear   INV 28-FEB-2001
DEFINITION  Drosophila melanogaster, chromosome 2R, region 49A-49B, BAC clone
            BACR04E21, complete sequence.
ACCESSION   AC007475
VERSION     AC007475.7  GI:13162478
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 173613)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome 2R, region 49A-49B
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 173613)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Feb 28, 2001 this sequence version replaced gi:6957978.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .173613
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="2R"
                     /map="49A-49B"
                     /clone="BACR04E21 (D592)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
BASE COUNT    48071 a  37844 c  38273 g  49425 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 3;  Length 173613;
  Best Local Similarity   90.0%;  Pred. No. 3.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||| ||||||||||  |||
Db      25802 CTGATGGCCAAACGAGAAAC 25821


RESULT 11
AC007474
LOCUS       AC007474              192763 bp    DNA     linear   INV 28-FEB-2001
DEFINITION  Drosophila melanogaster, chromosome 2R, region 49A-49B, BAC clone
            BACR14L19, complete sequence.
ACCESSION   AC007474
VERSION     AC007474.5  GI:13162477
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 192763)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome 2R, region 49A-49B
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 192763)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-MAY-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Feb 28, 2001 this sequence version replaced gi:5670617.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
            archive Web site (http://www.fruitfly.org/sequence/) or send email
            to bdgp@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .192763
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="2R"
```

```
                     /map="49A-49B"
                     /clone="BACR14L19 (D591)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
BASE COUNT     54505 a  41459 c  42141 g  54658 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 3;  Length 192763;
  Best Local Similarity   90.0%;  Pred. No. 3.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||| |||||||||| |||
Db     150079 CTGATGGCCAAACGAGAAAC 150098


RESULT 12
AC020286/c
LOCUS       AC020286               194634 bp    DNA     linear   HTG 03-JAN-2000
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AC020286
VERSION     AC020286.1  GI:6664611
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 194634)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10212663 by the submitter.
            For more information on this record e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .194634
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
BASE COUNT     55777 a  42412 c  41301 g  55144 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 2;  Length 194634;
  Best Local Similarity   90.0%;  Pred. No. 3.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||| ||||||||||| |||
Db      58864 CTGATGGCCAAACGAGAAAC 58845


RESULT 13
AC109049
LOCUS       AC109049               206732 bp    DNA     linear   HTG 09-OCT-2002
DEFINITION  Rattus norvegicus clone CH230-334A14, *** SEQUENCING IN PROGRESS
            ***, 2 unordered pieces.
ACCESSION   AC109049
VERSION     AC109049.4  GI:23101194
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 206732)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 206732)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-FEB-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 206732)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-OCT-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 18, 2002 this sequence version replaced gi:21737800.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
```

```
           table.
           -------------- Genome Center
               Center: Baylor College of Medicine
               Center code: BCM
               Web site: http://www.hgsc.bcm.tmc.edu/
               Contact: hgsc-help@bcm.tmc.edu
           -------------- Project Information
               Center project name: GPTJ
               Center clone name: CH230-334A14
           -------------- Summary Statistics
               Assembly program: Phrap; version 0.990329
               Consensus quality: 181594 bases at least Q40
               Consensus quality: 184576 bases at least Q30
               Consensus quality: 186655 bases at least Q20
               Estimated insert size: 197225; sum-of-contigs estimation
               Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
           --------------
           * NOTE: Estimated insert size may differ from sequence length
           *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 2 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *        1     3973: contig of 3973 bp in length
           *     3974     4073: gap of unknown length
           *     4074   206732: contig of 202659 bp in length.
FEATURES             Location/Qualifiers
     source          1. .206732
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-334A14"
     misc_feature    complement(484. .1138)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:MboI
                     end_sequence:RXAPN07TV"
     misc_feature    6011. .7297
                     /note="wgs_contig"
     misc_feature    complement(200587. .201453)
                     /note="clone_boundary
                     clone_end:T7
                     site:MboI
                     end_sequence:RXAPN07TJ"
BASE COUNT    54264 a  40169 c  41232 g  52474 t  18593 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 2;  Length 206732;
  Best Local Similarity   90.0%;  Pred. No. 3.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              |||||| |||||| ||||||
Db      65496 CTGATCACCAAACCAGCAAC 65515


RESULT 14
AC102969
LOCUS       AC102969               259345 bp    DNA     linear   HTG 10-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-94M19, *** SEQUENCING IN PROGRESS
            ***, 5 unordered pieces.
ACCESSION   AC102969
VERSION     AC102969.7  GI:30522593
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 259345)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 259345)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 259345)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:24819627.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
```

```
        contigs within a contig-scaffold that consist entirely of whole
        genome shotgun sequence reads. Both end sequences and whole genome
        shotgun sequence only contigs will be indicated in the feature
        table.
        -------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
        -------------- Project Information
            Center project name: GGOG
            Center clone name: CH230-94M19
        -------------- Summary Statistics
            Assembly program: Atlas 3.0;
            Consensus quality: 242269 bases at least Q40
            Consensus quality: 247151 bases at least Q30
            Consensus quality: 249910 bases at least Q20
            Estimated insert size: 260613; sum-of-contigs estimation
            Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
        --------------
        * NOTE: Estimated insert size may differ from sequence length
        *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
        * NOTE: This sequence may represent more than one clone.
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 5 contigs. The true order of the pieces
        * is not known and their order in this sequence record is
        * arbitrary. Gaps between the contigs are represented as
        * runs of N, but the exact sizes of the gaps are unknown.
        * This record will be updated with the finished sequence
        * as soon as it is available and the accession number will
        * be preserved.
        *       1   248466: contig of 248466 bp in length
        *  248467   248566: gap of unknown length
        *  248567   254507: contig of 5941 bp in length
        *  254508   254607: gap of unknown length
        *  254608   256455: contig of 1848 bp in length
        *  256456   256555: gap of unknown length
        *  256556   258160: contig of 1605 bp in length
        *  258161   258260: gap of unknown length
        *  258261   259345: contig of 1085 bp in length.
FEATURES             Location/Qualifiers
     source          1. .259345
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-94M19"
     misc_feature    1. .3071
                     /note="wgs_contig"
     misc_feature    complement(245699. .246677)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH355461"
     misc_feature    251214. .254507
                     /note="wgs_end_extension
                     clone_end:T7"
BASE COUNT    79986 a  53670 c  51512 g  66827 t   7350 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 2;  Length 259345;
  Best Local Similarity   90.0%;  Pred. No. 3.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              |||||| |||||| ||||||
Db     204951 CTGATCACCAAACCAGCAAC 204970


RESULT 15
AC103330
LOCUS       AC103330               261157 bp    DNA     linear   HTG 13-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-177K18, WORKING DRAFT SEQUENCE, 7
            unordered pieces.
ACCESSION   AC103330
VERSION     AC103330.5  GI:30580846
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 261157)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 261157)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 261157)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 13, 2003 this sequence version replaced gi:22857506.
```

```
          The sequence in this assembly is a combination of BAC based reads
          and whole genome shotgun sequencing reads assembled using Atlas
          (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
          in the feature table below represents a scaffold in the Atlas
          assembly (a 'contig-scaffold'). Within each contig-scaffold,
          individual sequence contigs are ordered and oriented, and separated
          by sized gaps filled with Ns to the estimated size. The sequence
          may extend beyond the ends of the clone and there may be sequence
          contigs within a contig-scaffold that consist entirely of whole
          genome shotgun sequence reads. Both end sequences and whole genome
          shotgun sequence only contigs will be indicated in the feature
          table.
          ---------- Genome Center
              Center: Baylor College of Medicine
              Center code: BCM
              Web site: http://www.hgsc.bcm.tmc.edu/
              Contact: hgsc-help@bcm.tmc.edu
          ---------- Project Information
              Center project name: GLAH
              Center clone name: CH230-177K18
          ---------- Summary Statistics
              Assembly program: Atlas 3.0;
              Consensus quality: 247708 bases at least Q40
              Consensus quality: 250022 bases at least Q30
              Consensus quality: 251823 bases at least Q20
              Estimated insert size: 259364; sum-of-contigs estimation
              Quality coverage: 9x in Q20 bases; sum-of-contigs estimation
          ----------
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
          * NOTE: This sequence may represent more than one clone.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 7 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *        1   242048: contig of 242048 bp in length
          *   242049   242148: gap of unknown length
          *   242149   246131: contig of 3983 bp in length
          *   246132   246231: gap of unknown length
          *   246232   252340: contig of 6109 bp in length
          *   252341   252440: gap of unknown length
          *   252441   253496: contig of 1056 bp in length
          *   253497   253596: gap of unknown length
          *   253597   255683: contig of 2087 bp in length
          *   255684   255783: gap of unknown length
          *   255784   257688: contig of 1905 bp in length
          *   257689   257788: gap of unknown length
          *   257789   261157: contig of 3369 bp in length.
FEATURES             Location/Qualifiers
     source          1. .261157
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-177K18"
     misc_feature    1. .2340
                     /note="wgs_contig"
     misc_feature    5033. .6658
                     /note="wgs_contig"
     misc_feature    239309. .242048
                     /note="wgs_contig"
     misc_feature    242149. .243506
                     /note="wgs_contig"
     misc_feature    246232. .247626
                     /note="wgs_contig"
BASE COUNT    66925 a  60485 c  60256 g  65490 t   8001 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 2;  Length 261157;
  Best Local Similarity   90.0%;  Pred. No. 3.5e+02;

  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              |||||||||| ||||||| |
Db     134133 CTGATCGCCAGACGAGCAGC 134152


Search completed: February 12, 2004, 04:41:47
Job time : 591.854 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 01:53:19 ; Search time 105.122 Seconds
(without alignments)
513.582 Million cell updates/sec

Title: US-09-692-077D-15
Perfect score: 20
Sequence: 1 ctgatcgccaaacgagcaac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	23	AAI99909	Human alpha-2BAR g
	2	16.8	84.0	3920	23	ABL22105	Drosophila melanog
c	3	16.8	84.0	6990	23	ABL22104	Drosophila melanog
c	4	15.8	79.0	277	25	ABX25006	Human GDP-mannose
	5	15.8	79.0	1976	25	ABX22622	Human GDP-mannose
	6	15.8	79.0	5786	22	AAS29976	Human lung antigen
	7	15.8	79.0	47999	25	AAD52898	Human tweety homol
	8	15.8	79.0	1082138	21	AAF22305	Arabidopsis thalia
c	9	15.4	77.0	543	24	ABN68569	Streptococcus poly
c	10	15.4	77.0	534720	19	AAV30458	Rhizobium species
c	11	15.4	77.0	536165	19	AAV30459	Rhizobium species
	12	15.2	76.0	144	24	ABQ91593	M. capsulatus gene
c	13	15.2	76.0	309	24	ABQ91594	M. capsulatus gene
c	14	15.2	76.0	705	24	ABQ65658	Arabidopsis thalia
	15	15.2	76.0	1281	25	ABT18455	Aspergillus fumiga
	16	15.2	76.0	1284	25	ABT19049	Aspergillus fumiga
c	17	15.2	76.0	1359	24	ABZ66820	Arabidopsis thalia
c	18	15.2	76.0	1359	24	ABZ14216	Arabidopsis thalia
c	19	15.2	76.0	1380	23	ABL14339	Drosophila melanog
c	20	15.2	76.0	1525	24	ABQ44152	Oligonucleotide fo
	21	15.2	76.0	1525	24	ABQ44153	Oligonucleotide fo
c	22	15.2	76.0	1525	24	ABQ45082	Oligonucleotide fo
	23	15.2	76.0	1525	24	ABQ45083	Oligonucleotide fo
c	24	15.2	76.0	1586	21	AAC45525	Arabidopsis thalia
c	25	15.2	76.0	1588	21	AAC41102	Arabidopsis thalia
	26	15.2	76.0	2036	23	ABL01847	Drosophila melanog
	27	15.2	76.0	2193	23	AAS53007	Enterococcus faeca
	28	15.2	76.0	2888	23	ABL02530	Drosophila melanog
c	29	15.2	76.0	3266	23	ABL01889	Drosophila melanog
	30	15.2	76.0	3281	25	ABT17861	Aspergillus fumiga
	31	15.2	76.0	3907	23	ABL14338	Drosophila melanog
c	32	15.2	76.0	4779	22	AAS29952	Human lung antigen
	33	15.2	76.0	6080	23	ABL01846	Drosophila melanog
c	34	15.2	76.0	7874	20	AAX20568	Polynucleotide seq
	35	15.2	76.0	8189	23	ABL01888	Drosophila melanog
c	36	15.2	76.0	1664976	19	AAV21209	Methanococcus jann
	37	15	75.0	1356	23	AAS94178	DNA encoding novel
	38	15	75.0	2416	23	ABL06847	Drosophila melanog
	39	15	75.0	2943	23	AAS86718	DNA encoding novel
	40	15	75.0	3681	23	ABL06346	Drosophila melanog
c	41	15	75.0	3963	23	AAS79164	DNA encoding novel
	42	15	75.0	5379	23	ABL06846	Drosophila melanog
	43	14.8	74.0	180	14	AAQ49414	Cytochrome P450 ho
c	44	14.8	74.0	304	21	AAC10249	Human secreted pro
c	45	14.8	74.0	391	25	ABX51029	Bovine EST associa

ALIGNMENTS

RESULT 1
AAI99909
ID AAI99909 standard; DNA; 20 BP.
XX
AC AAI99909;
XX
DT 18-FEB-2002 (first entry)
XX
DE Human alpha-2BAR genotyping PCR primer SEQ ID NO 15.
XX
KW Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW polymorphic site; allelic variant; cardiovascular disease;
KW central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW phosphorylation; inositol phosphate; alpha-2BAR; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200179561-A2.
XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;

XX
DR WPI; 2001-611728/70.
XX
PT Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT determining whether an individual is at increased risk of developing a
PT disease associated with the corresponding receptor comprises detecting
PT a polymorphic site -
XX
PS Claim 10; Page 112; 163pp; English.
XX
CC The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC alpha2A or alpha2C or fragment or complement of; and
CC (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC positions 961-972 of (III). The method may be used for genotyping an
CC alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC whether an individual is at increased risk of developing a disease
CC associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC polymorphic site which correlate to disease selected from cardiovascular
CC disease, central nervous system disease and combinations of these. In
CC addition, the technique may be used to predict an individual's response
CC to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC these) by detecting the polymorphic site and correlating the site to a
CC predetermined response (where the response is correlated to adenylyl
CC cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC levels). The present sequence is that of a human alpha-2BAR PCR primer,
CC useful for the genotyping methods of the invention.
XX
SQ Sequence 20 BP; 7 A; 7 C; 4 G; 2 T; 0 other;

Query Match 100.0%; Score 20; DB 23; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||||||||||||||||||
Db          1 CTGATCGCCAAACGAGCAAC 20
```

RESULT 2
ABL22105
ID ABL22105 standard; DNA; 3920 BP.
XX
AC ABL22105;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 17788.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 17788; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3920 BP; 1053 A; 1007 C; 947 G; 913 T; 0 other;

Query Match 84.0%; Score 16.8; DB 23; Length 3920;
Best Local Similarity 90.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||| |||||||||| |||
Db       3432 CTGATGGCCAAACGAGAAAC 3451
```

RESULT 3
ABL22104/c
ID ABL22104 standard; DNA; 6990 BP.
XX
AC ABL22104;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 17785.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 17785; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

```
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 6990 BP; 1907 A; 1513 C; 1586 G; 1984 T; 0 other;

  Query Match             84.0%;  Score 16.8;  DB 23;  Length 6990;
  Best Local Similarity   90.0%;  Pred. No. 60;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||| |||||||||| |||
Db       1624 CTGATGGCCAAACGAGAAAC 1605


RESULT 4
ABX25006/c
ID   ABX25006 standard; cDNA; 277 BP.
XX
AC   ABX25006;
XX
DT   11-FEB-2003' (first entry)
XX
DE   Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #7063.
XX
KW   Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
KW   cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW   arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW   complex carbohydrate; gene replacement therapy; immunosuppressive;
KW   antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW   antiasthmatic; vasotropic.
XX
OS   Homo sapiens.
XX
PN   US2002110548-A1.
XX
PD   15-AUG-2002.
XX
PF   11-JUN-2001; 2001US-0878574.
XX
PR   22-NOV-1996;   96US-0753233.
PR   03-DEC-1997;   97US-0984246.
PR   09-SEP-1998;   98US-0149674.
PR   14-JUN-1999;   99US-0333177.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Sullivan F,  Kriz R,  Kumar R;
XX
DR   WPI; 2003-066673/06.
XX
PT   New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT   peptide, for manufacturing complex carbohydrates, or as targets for
PT   screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT   rejection  -
XX
PS   Disclosure; SEQ ID NO 7065; 6pp; English.
XX
CC   The invention relates to a composition comprising a human GDP-mannose
CC   4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
CC   GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
CC   in a mammalian subject and for treating or ameliorating diseases affected
CC   by the level of cellular fucosylation or diseases affected by the
CC   fucosylation of glycoconjugates. These diseases include arthritis,
CC   transplant rejection, asthma, sepsis, reperfusion injury, stroke or
CC   infection. The GM4,6D peptide or a polynucleotide encoding it is also
CC   useful for manufacturing complex carbohydrates and as targets for
CC   screening small molecule antagonists of the activity of the enzyme. The
CC   polynucleotide is useful in developing an assay for defects in the
CC   enzyme, as well as in gene replacement therapy. Sequences
CC   ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
CC   human GM4,6D peptides of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 277 BP; 89 A; 57 C; 64 G; 67 T; 0 other;

  Query Match             79.0%;  Score 15.8;  DB 25;  Length 277;
  Best Local Similarity   89.5%;  Pred. No. 1.5e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAAC 20
              |||||||||| ||| ||||
Db        277 TGATCGCCAAGCGAACAAC 259


RESULT 5
ABX22622
ID   ABX22622 standard; cDNA; 1976 BP.
XX
AC   ABX22622;
XX
DT   10-FEB-2003  (first entry)
XX
DE   Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #4679.
XX
KW   Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
KW   cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
KW   arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
KW   complex carbohydrate; gene replacement therapy; immunosuppressive;
KW   antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
KW   antiasthmatic; vasotropic.
XX
OS   Homo sapiens.
XX
PN   US2002110548-A1.
XX
PD   15-AUG-2002.
XX
PF   11-JUN-2001; 2001US-0878574.
XX
PR   22-NOV-1996;   96US-0753233.
PR   03-DEC-1997;   97US-0984246.
PR   09-SEP-1998;   98US-0149674.
PR   14-JUN-1999;   99US-0333177.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Sullivan F,  Kriz R,  Kumar R;
XX
DR   WPI; 2003-066673/06.
XX
PT   New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT   peptide, for manufacturing complex carbohydrates, or as targets for
PT   screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT   rejection  -
XX
PS   Disclosure; SEQ ID NO 4681; 6pp; English.
XX
CC   The invention relates to a composition comprising a human GDP-mannose
CC   4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
CC   GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
CC   in a mammalian subject and for treating or ameliorating diseases affected
CC   by the level of cellular fucosylation or diseases affected by the
CC   fucosylation of glycoconjugates. These diseases include arthritis,
CC   transplant rejection, asthma, sepsis, reperfusion injury, stroke or
CC   infection. The GM4,6D peptide or a polynucleotide encoding it is also
CC   useful for manufacturing complex carbohydrates and as targets for
CC   screening small molecule antagonists of the activity of the enzyme. The
CC   polynucleotide is useful in developing an assay for defects in the
CC   enzyme, as well as in gene replacement therapy. Sequences
CC   ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
CC   human GM4,6D peptides of the invention.
```

```
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 1976 BP; 497 A; 477 C; 374 G; 522 T; 106 other;

  Query Match             79.0%;  Score 15.8;  DB 25;  Length 1976;
  Best Local Similarity   85.0%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||||| ||| ||| ||||
Db       1174 CTGATCGACAANCGATCAAC 1193


RESULT 6
AAS29976
ID   AAS29976 standard; DNA; 5786 BP.
XX
AC   AAS29976;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human lung antigen genomic DNA #46.
XX
KW   Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;
KW   chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW   cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW   ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW   hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW   cerebrovascular disorder; nervous system disorder; bacterial infection;
KW   fungal infection; viral infection; ocular disorder; endocrine disorder;
KW   gastrointestinal disorder; renal disorder; respiratory disorder;
KW   wound healing; skin aging; organ transplantation; food preservative;
KW   tissue regeneration; anti-infertility; food additive.
XX
OS   Homo sapiens.
XX
PN   WO200155303-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01301.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
```

```
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-457723/49.
XX
PT   Isolated polypeptide for treating, preventing and/or prognosing
PT   respiratory disorders related to the lung including lung cancers and
PT   also for testing and detection e.g. diagnosis -
XX
PS   Claim 1; SEQ ID No 240; 507pp; English.
XX
CC   Sequences AAS29931-AAS30164 represent genomic DNA molecules, which encode
CC   the lung antigen polypeptides of the invention. Lung antigen polypeptides
CC   and their associated polynucleotides are useful in the diagnosis,
CC   treatment and prevention of various types of disorders in e.g. humans,
CC   mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC   pathological condition can be determined by detecting the presence or
CC   absence of a mutation in a lung antigen polynucleotide. The treatable
CC   disorders include autoimmune diseases such as rheumatoid arthritis,
CC   hyperproliferative disorders such as neoplasms of the breast or liver,
CC   cardiovascular disorders such as cardiac arrest, cerebrovascular
CC   disorders such as cerebral ischaemia, nervous system disorders such as
CC   Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC   ocular disorders such as corneal infection, endocrine disorders such as
CC   premature labour and infertility, gastrointestinal disorders such as
CC   Crohn's disease, renal disorders such as glomerulonephritis and
CC   respiratory disorders such as asthma and pleurisy. The polypeptides can
CC   also be used to aid wound healing, to prevent skin aging due to sunburn,
CC   to maintain organs before transplantation, to regenerate tissues and in
CC   chemotaxis. The polypeptides can also be used as a food additive or
CC   preservative to increase or decrease storage capabilities.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.

  Query Match             79.0%;  Score 15.8;  DB 22;  Length 5786;
  Best Local Similarity   89.5%;  Pred. No. 1.9e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAAC 20
              | ||||||||| |||||||
Db        818 TTATCGCCAAATGAGCAAC 836


RESULT 7
AAD52898
ID   AAD52898 standard; DNA; 47999 BP.
XX
AC   AAD52898;
XX
DT   14-MAY-2003  (first entry)
XX
DE   Human tweety homologue 2 (TTYH2) gene.
XX
KW   Human; tweety homologue 2; TTYH2; therapy; cancer; tumour; cytostatic;
KW   diagnostic marker; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1946..45753
FT                   /*tag=  a
FT                   /product= "Human TTYH2 protein"
FT   exon            1936..2074
FT                   /*tag=  b
FT                   /number= 1
FT   intron          2075..10376
FT                   /*tag=  c
FT   exon            10377..10549
FT                   /*tag=  d
FT                   /number= 2
FT   intron          10550..16622
FT                   /*tag=  e
FT   exon            16623..16734
FT                   /*tag=  f
FT                   /number= 3
FT   intron          16735..23223
FT                   /*tag=  g
FT   exon            23224..23444
FT                   /*tag=  h
FT                   /number= 4
FT   intron          23445..28299
FT                   /*tag=  i
FT   exon            28300..28395
FT                   /*tag=  j
FT                   /number= 5
FT   intron          28396..28902
FT                   /*tag=  k
FT   exon            28903..28975
FT                   /*tag=  l
FT                   /number= 6
FT   intron          28976..35372
FT                   /*tag=  m
FT   exon            35373..35442
FT                   /*tag=  n
FT                   /number= 7
FT   intron          35443..35705
FT                   /*tag=  o
FT   exon            35706..35761
FT                   /*tag=  p
FT                   /number= 8
FT   intron          35762..36266
FT                   /*tag=  q
FT   exon            36267..36359
FT                   /*tag=  r
FT                   /number= 9
FT   intron          36360..36591
FT                   /*tag=  s
FT   exon            36592..36684
```

```
FT                   /*tag=  t
FT                   /number= 10
FT   intron          36685..38529
FT                   /*tag=  u
FT   exon            38530..38672
FT                   /*tag=  v
FT                   /number= 11
FT   intron          38673..39376
FT                   /*tag=  w
FT   exon            39377..39562
FT                   /*tag=  x
FT                   /number= 12
FT   intron          39563..40050
FT                   /*tag=  y
FT   exon            40051..40129
FT                   /*tag=  z
FT                   /number= 13
FT   intron          40130..45672
FT                   /*tag=  aa
FT   exon            45673..47158
FT                   /*tag=  ab
FT                   /number= 14
XX
PN   WO200292629-A1.
XX
PD   21-NOV-2002.
XX
PF   14-MAY-2002; 2002WO-AU00591.
XX
PR   14-MAY-2001; 2001AU-0004971.
XX
PA   (UYQU-) UNIV QUEENSLAND TECHNOLOGY.
XX
PI   Clements JA;
XX
DR   WPI; 2003-129264/12.
DR   P-PSDB; AAE34613.
XX
PT   New human tweety homolog 2 polypeptides and polynucleotides, useful for
PT   producing an antigen-binding molecule that is immuno-interactive with
PT   the polypeptide or as diagnostic markers for cancers -
XX
PS   Claim 10; Page 128-156; 176pp; English.
XX
CC   The invention relates to human tweety homologue 2 (TTYH2) polypeptide and
CC   polynucleotide sequence. TTYH2 is useful for producing an antigen-binding
CC   molecule that is immuno-interactive with the polypeptide. The agent is
CC   useful for manufacturing a medicament for restoring a normal level and/or
CC   functional activity of TTYH2 expression in a patient, and for treating or
CC   preventing cancer or tumour. TTYH2 sequences may also be used to provide
CC   both drug targets and regulators to promote or inhibit one or more
CC   activities, and to provide diagnostic markers for cancers. The present
CC   sequence is human TTYH2 gene.
XX
SQ   Sequence 47999 BP; 11178 A; 12727 C; 12874 G; 11209 T; 11 other;

  Query Match             79.0%;  Score 15.8;  DB 25;  Length 47999;
  Best Local Similarity   89.5%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAA 19
              ||| || ||||||||||||
Db      28575 CTGGTCACCAAACGAGCAA 28593


RESULT 8
AAF22305
ID   AAF22305 standard; DNA; 1082138 BP.
XX
AC   AAF22305;
XX
DT   20-MAR-2001  (first entry)
XX
DE   Arabidopsis thaliana chromosome 4 centromere.
XX
KW   Centromere; michrosome; vector; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200055325-A2.
XX
PD   21-SEP-2000.
XX
PF   17-MAR-2000; 2000WO-US07392.
XX
PR   18-MAR-1999;   99US-0125219.
PR   01-APR-1999;   99US-0127409.
PR   18-MAY-1999;   99US-0134770.
PR   13-SEP-1999;   99US-0153584.
PR   17-SEP-1999;   99US-0154603.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Preuss D,  Copenhaver G,  Keith K;
XX
DR   WPI; 2000-587529/55.
XX
PT   Recombinant DNA construct comprising a plant centromere, useful for
PT   producing stably inherited michrosomes which can serve as vectors for
PT   the construction of transgenic plant and animal cells -
XX
PS   Claim 68; Page 977-1388; 1449pp; English.
XX
CC   The present invention relates to a recombinant DNA construct of a plant
CC   (Arabidopsis thaliana) centromere. The constructs are useful for
CC   producing stably inherited michrosomes which can serve as vectors for
CC   the construction of transgenic plant and animal cells expressing
CC   selected proteins such as hormones, enzymes, interleukins, clotting
CC   factors, cytokines, antibodies, and growth factors.
XX
SQ   Sequence 1082138 BP; 348775 A; 194404 C; 195515 G; 343444 T; 0 other;

  Query Match             79.0%;  Score 15.8;  DB 21;  Length 1082138;
  Best Local Similarity   89.5%;  Pred. No. 2.8e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAAC 20
              || ||||||| ||||||||
Db     734777 TGTTCGCCAAGCGAGCAAC 734795


RESULT 9
ABN68569/c
ID   ABN68569 standard; DNA; 543 BP.
XX
AC   ABN68569;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Streptococcus polynucleotide SEQ ID NO 5051.
XX
KW   Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW   group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW   antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS   Streptococcus pyogenes.
XX
PN   WO200234771-A2.
XX
PD   02-MAY-2002.
XX
PF   29-OCT-2001; 2001WO-GB04789.
XX
PR   27-OCT-2000; 2000GB-0026333.
PR   24-NOV-2000; 2000GB-0028727.
PR   07-MAR-2001; 2001GB-0005640.
```

```
XX
PA   (CHIR-) CHIRON SPA.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Telford J,  Masignani V,  Margarit Ros YI,  Grandi G,  Fraser C;
PI   Tettelin H;
XX
DR   WPI; 2002-352536/38.
DR   P-PSDB; ABP27938.
XX
PT   New Streptococcus protein for the treatment or prevention of infection
PT   or disease caused by Streptococcus bacteria, such as meningitis, and
PT   for detecting a compound that binds to the protein -
XX
PS   Claim 7; Page 3667; 4525pp; English.
XX
CC   The invention relates to a protein (ABP25413-ABP30895) from group B
CC   streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC   (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC   the specification. The proteins have antibacterial and antiinflammatory
CC   activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC   antibodies that bind (I) are used in the manufacture of medicaments for
CC   the treatment or prevention of infection or disease caused by
CC   Streptococcus bacteria, particularly S. agalactiae and S. pyrogenes.
CC   Nucleic acids encoding (I) are used to detect Streptococcus in a
CC   biological sample. (I) is used to determine whether a compound binds to
CC   (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC   used as a vaccine or diagnostic composition. The disease caused by
CC   Streptococcus that is prevented or treated may be meningitis. Nucleic
CC   acid encoding (I) may be used to recombinantly produce (I) and may be
CC   used in gene therapy. Antibodies to (I) are used for affinity
CC   chromatography, immunoassays, and distinguishing/identifying
CC   Streptococcus proteins.
XX
SQ   Sequence 543 BP; 123 A; 90 C; 93 G; 237 T; 0 other;

  Query Match             77.0%;  Score 15.4;  DB 24;  Length 543;
  Best Local Similarity   94.1%;  Pred. No. 2.6e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GATCGCCAAACGAGCAA 19
              |||| ||||||||||||
Db        486 GATCACCAAACGAGCAA 470


RESULT 10
AAV30458/c
ID   AAV30458 standard; DNA; 534720 BP.
XX
AC   AAV30458;
XX
DT   14-OCT-1998  (first entry)
XX
DE   Rhizobium species plasmid pNGR234a.
XX
KW   Symbiosis; open reading frame; ORF; plasmid; vector; transportation;
KW   degradation; metabolism; host range; nitrogen fixation; nodulation;
KW   legume; plant; ds.
XX
OS   Rhizobium sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             417796..418671
FT                   /*tag=  a
FT                   /standard_name= "ORF K1"
FT                   /product= "oligopeptide permease"
FT                   /note= "homologous to the OppC gene"
FT   CDS             418673..419680
FT                   /*tag=  b
FT                   /standard_name= "ORF K2"
FT                   /product= "oligopeptide permease"
FT                   /note= "homologous to the OppD gene"
FT   CDS             419677..420738
FT                   /*tag=  c
FT                   /standard_name= "ORF K3"
FT                   /product= "oligopeptide permease"
FT                   /note= "homologous to the OppF gene"
FT   CDS             420774..422159
FT                   /*tag=  d
FT                   /standard_name= "ORF K4"
FT                   /product= "encapsulation-like protein"
FT                   /note= "homologous to the CapA gene"
FT   CDS             422628..424031
FT                   /*tag=  e
FT                   /standard_name= "ORF K5"
FT                   /product= "aminotransferase-like protein"
FT                   /note= "homologous to the BioA gene"
FT   CDS             424056..425594
FT                   /*tag=  f
FT                   /standard_name= "ORF K6"
FT                   /product= "(semi)aldehyde dehydrogenase-like protein"
FT   CDS             complement (426949..428028)
FT                   /*tag=  g
FT                   /standard_name= "ORF K7"
FT                   /product= "transposase homologue"
FT                   /note= "homologous to the Tnp gene"
FT   CDS             428292..429623
FT                   /*tag=  h
FT                   /standard_name= "ORF K8"
FT                   /product= "glutamate dehydrogenase-like protein"
FT                   /note= "homologous to the GLUD1 gene"
FT   CDS             complement (430538..431284)
FT                   /*tag=  i
FT                   /standard_name= "ORF K9"
FT                   /product= "transposase homologue"
FT   CDS             complement (431296..432840)
FT                   /*tag=  j
FT                   /standard_name= "ORF K10"
FT                   /product= "transposase homologue"
FT                   /note= "homologous to the Tnp gene"
FT   CDS             complement (433880..434110)
FT                   /*tag=  k
FT                   /standard_name= "ORF K11"
FT                   /product= "protein of unknown function"
FT                   /note= "homologous to the FixU gene"
FT   CDS             complement (434107..434433)
FT                   /*tag=  l
FT                   /standard_name= "ORF K12"
FT                   /product= "protein of unknown function"
FT   CDS             complement (434517..434711)
FT                   /*tag=  m
FT                   /standard_name= "ORF K13"
FT                   /product= "ferrodoxin/ferrodoxin-like protein"
FT                   /note= "homologous to the FdxN gene"
FT   CDS             complement (434753..436234)
FT                   /*tag=  n
FT                   /standard_name= "ORF K14"
FT                   /gene= "nifB"
FT                   /product= "protein involved in FeMo co-factor
FT                            biosynthesis"
FT   CDS             complement (436460..438130)
FT                   /*tag=  o
FT                   /standard_name= "ORF K15"
FT                   /gene= "nifA"
FT                   /product= "positive regulator of nif, fix and other
FT                            genes"
FT   CDS             complement (438297..438590)
FT                   /*tag=  p
FT                   /standard_name= "ORF K16"
FT                   /gene= "fixX"
FT                   /product= "protein required for nitrogenase activity"
FT   CDS             complement (438605..439912)
FT                   /*tag=  q
FT                   /standard_name= "ORF K17"
FT                   /gene= "fixC"
FT                   /product= "protein required for nitrogenase activity"
```

```
FT   CDS             complement (439923..441032)
FT                   /*tag=  r
FT                   /standard_name= "ORF K18"
FT                   /gene= "fixB"
FT                   /product= "protein required for nitrogenase activity"
FT   CDS             complement (441042..441899)
FT                   /*tag=  s
FT                   /standard_name= "ORF K19"
FT                   /gene= "fixA"
FT                   /product= "protein required for nitrogenase activity"
FT   CDS             complement (442316..442636)
FT                   /*tag=  t
FT                   /standard_name= "ORF K20
FT                   /product= "protein of unknown function"
FT   CDS             complement (443313..443879)
FT                   /*tag=  u
FT                   /standard_name= "ORF K21"
FT                   /product= "protein of unknown function"
FT   CDS             444337..445029
FT                   /*tag=  v
FT                   /standard_name= "ORF K22"
FT                   /product= "ferrodoxin-like protein"
FT                   /note= "homologous to the NifQ gene"
FT   CDS             445088..446602
FT                   /*tag=  w
FT                   /standard_name= "ORF K23"
FT                   /gene= "dctA"
FT                   /product= "C4-dicarboxylate transport protein"
FT                   /note= "homologous to the DctAI gene"
FT   CDS             446599..447843
FT                   /*tag=  x
FT                   /standard_name= "ORF L1"
FT                   /product= "cytochrome P450-like protein"
FT                   /note= "homologous to the CamC gene"
FT   CDS             447844..448500
FT                   /*tag=  y
FT                   /standard_name= "ORF L2"
FT                   /product= "gamma-hexachlorocyclohexane-dechlorinase-like
FT                            protein"
FT                   /note= "homologous to the LinA gene"
FT   CDS             448497..450203
FT                   /*tag=  z
FT                   /standard_name= "ORF L3"
FT                   /product= "putative protein with degradative function"
FT   CDS             450341..451396
FT                   /*tag=  aa
FT                   /standard_name= "ORF L4"
FT                   /product= "luciferase alpha-subunit-like protein"
FT                   /note= "homologous to the LuxA gene"
FT   CDS             452980..454494
FT                   /*tag=  ab
FT                   /standard_name= "ORF L6"
FT                   /gene= "nifD"
FT                   /product= "alpha-subunit of FeMo protein of nitrogenase"
FT   CDS             454590..456131
FT                   /*tag=  ac
FT                   /standard_name= "ORF L7"
FT                   /gene= "nifK"
FT                   /product= "beta-subunit of FeMo protein of nitrogenase"
FT   CDS             456187..457677
FT                   /*tag=  ad
FT                   /standard_name= "ORF L8"
FT                   /product= "protein involved in FeMo co-factor
FT                            biosynthesis"
FT                   /note= "homologous to the NifB gene"
FT   CDS             457687..459096
FT                   /*tag=  ae
FT                   /standard_name= "ORF L9"
FT                   /product= "protein involved in FeMo co-factor
FT                            biosynthesis"
FT                   /note= "homologous to the FixF gene"
FT   CDS             459093..459575
FT                   /*tag=  af
FT                   /standard_name= "ORF L10"
FT                   /product= "protein of unknown function"
FT                   /note= "homologous to the NifX gene"
FT   CDS             459579..460067
FT                   /*tag=  ag
FT                   /standard_name= "ORF L11"
FT                   /product= "protein of unknown function"
FT   CDS             460501..460920
FT                   /*tag=  ah
FT                   /standard_name= "ORF L12"
FT                   /product= "protein similar to part of the Fe protein
FT                            of nitrogenase"
FT                   /note= "homologous to the NifH gene"
FT   CDS             461228..461545
FT                   /*tag=  ai
FT                   /standard_name= "ORF L13"
FT                   /product= "protein of unknown function"
FT   CDS             463201..464739
FT                   /*tag=  aj
FT                   /standard_name= "ORF L14"
FT                   /product= "peptidase-like protein"
FT                   /note= "homologous to the bI-MPP gene"
FT   CDS             464736..466079
FT                   /*tag=  ak
FT                   /standard_name= "ORF L15"
FT                   /product= "processing protease-like protein"
FT                   /note= "homologous to the PP gene"
FT   CDS             466590..467021
```

Query Match 77.0%; Score 15.4; DB 19; Length 534720;
Best Local Similarity 94.1%; Pred. No. 4.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 TGATCGCCAAACGAGCA 18
              |||||||||| ||||||
Db     433213 TGATCGCCAATCGAGCA 433197
```

```
RESULT 11
AAV30459/c
ID   AAV30459 standard; DNA; 536165 BP.
XX
AC   AAV30459;
XX
DT   06-JUL-1999  (first entry)
XX
DE   Rhizobium species symbiotic plasmid pNGR234.
XX
KW   Symbiosis; open reading frame; ORF; plasmid; vector; transportation;
KW   degradation; metabolism; host range; nitrogen fixation; nodulation;
KW   legume; plant; ds.
XX
OS   Rhizobium sp.
XX
PN   WO9802560-A2.
XX
PD   22-JAN-1998.
XX
PF   10-JUL-1997;   97WO-IB00950.
XX
PR   20-MAY-1997;   97GB-0010395.
PR   12-JUL-1996;   96EP-0730001.
XX
PA   (MOLE-) INST MOLECULAR BIOTECHNOLOGY.
PA   (BIOL-) LAB BIOLOGIE MOLECULAIRE PLANTES SUPERIE.
XX
PI   Broughton WJ,  Freiberg CB,  Perret XP,  Rosenthal A;
XX
DR   WPI; 1998-110606/10.
XX
PT   New isolated symbiotic plasmid from Rhizobium sp. NGR234 - used to
PT   develop products for modifying plant characteristics, e.g. nitrogen
PT   fixation, synthesis of compounds and stress response
```

```
XX
PS   Claim 1; Fig 3; 228pp; English.
XX
CC   This is the nucleotide sequence of the plasmid pNGR234a isolated from
CC   Rhizobium sp. NGR234.  Open reading frames (ORF) derivable from the
CC   nucleotide sequence are claimed.  The nucleotide sequences or ORFs can
CC   be used e.g. in the transportation of compounds to and from an organism
CC   which is a host to at least one of the nucleotide sequences, ORFs or
CC   proteins, the degradation and/or metabolism of organic, inorganic,
CC   natural or xenobiotic substances in a host organism or the modification
CC   of the host range, nitrogen fixation abilities; for obtaining a synthetic
CC   minimal set of ORFs required for functional Rhizobium-legume symbiosis,
CC   especially for nodulation efficiency on host plants.
XX
SQ   Sequence 536165 BP; 111291 A; 155755 C; 157864 G; 111255 T; 0 other;

  Query Match             77.0%;  Score 15.4;  DB 19;  Length 536165;
  Best Local Similarity   94.1%;  Pred. No. 4.3e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCA 18
              ||||||||||  ||||||
Db     433213 TGATCGCCAATCGAGCA 433197


RESULT 12
ABQ91593
ID   ABQ91593 standard; DNA; 144 BP.
XX
AC   ABQ91593;
XX
DT   01-OCT-2002  (first entry)
XX
DE   M. capsulatus gene #1578 for DNA array.
XX
KW   Micro array; gene; ds; differential expression; gene expression.
XX
OS   Methylococcus capsulatus.
XX
PN   WO200255655-A2.
XX
PD   18-JUL-2002.
XX
PF   14-JAN-2002; 2002WO-NO00019.
XX
PR   12-JAN-2001; 2001NO-0000235.
PR   12-JAN-2001; 2001NO-0000239.
XX
PA   (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
PA   (TIGR-) TIGR.
XX
PI   Birkeland NK,  Eidhammer I,  Jonassen I,  Jensen HB,  Lien T;
PI   Lillehaug JR,  Lossius I,  Eisen JA,  Fraser CM,  Durkin AS;
PI   Salzberg SL;
XX
DR   WPI; 2002-557818/59.
XX
PT   Novel DNA array useful for determining differential expression of
PT   Methylococcus capsulatus genes, comprises polynucleotides or
PT   oligonucleotides representative for a selective number of Methylococcus
PT   capsulatus genes  -
XX
PS   Claim 14; Page 604; 678pp; English.
XX
CC   The invention relates to a novel DNA array giving a representation of a
CC   number of Methylococcus capsulatus genes. The method of the invention is
CC   useful for determination of the differential expression of the genes of
CC   M. capsulatus, and for studying gene expression on a genomic scale and in
CC   gene expression assays of M. capsulatus genes. The sequences shown in
CC   ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the
CC   invention.
XX
SQ   Sequence 144 BP; 30 A; 50 C; 33 G; 31 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 24;  Length 144;
  Best Local Similarity   85.0%;  Pred. No. 2.9e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||||||||  |||  ||  ||
Db         97 CTGATCGCCATACGGGCGAC 116


RESULT 13
ABQ91594/c
ID   ABQ91594 standard; DNA; 309 BP.
XX
AC   ABQ91594;
XX
DT   01-OCT-2002  (first entry)
XX
DE   M. capsulatus gene #1579 for DNA array.
XX
KW   Micro array; gene; ds; differential expression; gene expression.
XX
OS   Methylococcus capsulatus.
XX
PN   WO200255655-A2.
XX
PD   18-JUL-2002.
XX
PF   14-JAN-2002; 2002WO-NO00019.
XX
PR   12-JAN-2001; 2001NO-0000235.
PR   12-JAN-2001; 2001NO-0000239.
XX
PA   (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
PA   (TIGR-) TIGR.
XX
PI   Birkeland NK,  Eidhammer I,  Jonassen I,  Jensen HB,  Lien T;
PI   Lillehaug JR,  Lossius I,  Eisen JA,  Fraser CM,  Durkin AS;
PI   Salzberg SL;
XX
DR   WPI; 2002-557818/59.
XX
PT   Novel DNA array useful for determining differential expression of
PT   Methylococcus capsulatus genes, comprises polynucleotides or
PT   oligonucleotides representative for a selective number of Methylococcus
PT   capsulatus genes  -
XX
PS   Claim 14; Page 604; 678pp; English.
XX
CC   The invention relates to a novel DNA array giving a representation of a
CC   number of Methylococcus capsulatus genes. The method of the invention is
CC   useful for determination of the differential expression of the genes of
CC   M. capsulatus, and for studying gene expression on a genomic scale and in
CC   gene expression assays of M. capsulatus genes. The sequences shown in
CC   ABQ90016-ABQ91855 represent M. capsulatus genes for use in arrays of the
CC   invention.
XX
SQ   Sequence 309 BP; 51 A; 86 C; 109 G; 63 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 24;  Length 309;
  Best Local Similarity   85.0%;  Pred. No. 3.1e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||||||||  |||  ||  ||
Db        125 CTGATCGCCATACGGGCGAC 106


RESULT 14
ABQ65658/c
ID   ABQ65658 standard; DNA; 705 BP.
XX
AC   ABQ65658;
```

XX
DT 21-AUG-2002 (first entry)
XX
DE Arabidopsis thaliana polynucleotide SEQ ID NO 235.
XX
KW Arabidopsis thaliana; thale cress; plant; transgenic; GMO; disease;
KW stress; metabolic pathway; biosynthetic pathway; nutrition; fungicide;
KW insecticide; antibiotic; ds.
XX
OS Arabidopsis thaliana.
XX
PN US2002059663-A1.
XX
PD 16-MAY-2002.
XX
PF 26-JAN-2001; 2001US-0770149.
XX
PR 27-JAN-2000; 2000US-178506P.
XX
PA (GORL/) GORLACH J.
PA (ANYY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUYY/) YU Y.
PA (RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHEW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
DR WPI; 2002-479224/51.
XX
PT New nucleic acid that hybridizes to Arabidopsis thaliana sequences,
PT useful e.g. for preparing transgenic plants with increased resistance
PT or altered metabolism -
XX
PS Claim 1; SEQ ID NO 235; 40pp + Sequence Listing; English.
XX
CC The invention relates to nucleic acids (I) that hybridise under stringent
CC conditions to any of 999 sequences (ABQ65424-ABQ66422) or their
CC fragments. (I) are used to express the corresponding polypeptides (II) or
CC to produce genetically modified plant cells or transgenic plants, which
CC may have improved resistance to disease or stress, or altered
CC metabolic/biosynthetic pathways (for production of commercial,
CC nutritional or medicinal products), or generally any trait of interest,
CC or can be used to screen for biologically active agents (e.g. fungicides,
CC insecticides and antibiotics).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO at seqdata.uspto.gov/sequence.html?DocID=999909770149.
XX
SQ Sequence 705 BP; 183 A; 135 C; 183 G; 200 T; 4 other;

Query Match 76.0%; Score 15.2; DB 24; Length 705;
Best Local Similarity 85.0%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTGATCGCCAAACGAGCAAC 20
|| ||||||| ||||||| |
Db 182 CTCATCGCCAGACGAGCACC 163

RESULT 15
ABT18455
ID ABT18455 standard; DNA; 1281 BP.
XX
AC ABT18455;
XX
DT 16-APR-2003 (first entry)
XX
DE Aspergillus fumigatus essential gene #813.
XX
KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
KW cancer; contamination; biofilm; antibody; immune response; ds.
XX
OS Aspergillus fumigatus.
XX
PN WO200286090-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US13142.
XX
PR 23-APR-2001; 2001US-285697P.
PR 27-APR-2001; 2001US-287066P.
PR 05-JUN-2001; 2001US-295890P.
PR 09-JUL-2001; 2001US-303899P.
PR 31-AUG-2001; 2001US-316362P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX
DR WPI; 2003-093124/08.
XX
PT New purified or isolated nucleic acids of essential genes of
PT Aspergillus fumigatus, useful for treating or preventing infections by
PT A. fumigatus, or for treating a non-infectious disease in a subject
PT e.g. cancer -
XX
PS Disclosure; Page -; 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as A. fumigatus, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC biofilm comprising A. fumigatus. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of A.
CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case or virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of Aspergillus fumigatus of the invention.
XX
SQ Sequence 1281 BP; 300 A; 358 C; 333 G; 290 T; 0 other;

Query Match 76.0%; Score 15.2; DB 25; Length 1281;

Best Local Similarity 85.0%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||| |||||||||| || ||
Db         37 CTGGTCGCCAAACGCGCCAC 56
```

Search completed: February 12, 2004, 04:01:31
Job time : 109.122 secs

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 04:01:41 ; Search time 27.3171 Seconds
(without alignments)
323.156 Million cell updates/sec

Title: US-09-692-077D-15
Perfect score: 20
Sequence: 1 ctgatcgccaaacgagcaac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15.4	77.0	1698	4	US-09-328-352-643	Sequence 643, App
	2	15.4	77.0	3032	4	US-09-221-017B-347	Sequence 347, App
c	3	15.4	77.0	536165	4	US-09-214-808-1	Sequence 1, Appli
	4	15.2	76.0	36470	4	US-08-311-731A-123	Sequence 123, App
c	5	15.2	76.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	6	14.8	74.0	543	1	US-07-912-900-25	Sequence 25, Appl
	7	14.8	74.0	543	1	US-08-285-309-25	Sequence 25, Appl
	8	14.8	74.0	543	2	US-08-502-046-25	Sequence 25, Appl
	9	14.8	74.0	547	1	US-08-313-075A-31	Sequence 31, Appl
c	10	14.8	74.0	810	4	US-09-107-532A-2028	Sequence 2028, Ap
	11	14.8	74.0	4379	1	US-08-592-214A-17	Sequence 17, Appl
	12	14.8	74.0	4379	3	US-09-149-976-17	Sequence 17, Appl
	13	14.8	74.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	14	14.8	74.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	15	14.2	71.0	405	4	US-09-643-597-218	Sequence 218, App
	16	14.2	71.0	405	4	US-09-480-884A-218	Sequence 218, App
	17	14.2	71.0	405	4	US-09-542-615A-218	Sequence 218, App
	18	14.2	71.0	405	4	US-09-606-421B-218	Sequence 218, App
	19	14.2	71.0	423	3	US-09-040-984-60	Sequence 60, Appl
	20	14.2	71.0	423	4	US-09-123-912-60	Sequence 60, Appl
	21	14.2	71.0	423	4	US-09-643-597-60	Sequence 60, Appl
	22	14.2	71.0	423	4	US-09-480-884A-60	Sequence 60, Appl
	23	14.2	71.0	423	4	US-09-542-615A-60	Sequence 60, Appl
	24	14.2	71.0	423	4	US-09-606-421B-60	Sequence 60, Appl
c	25	14.2	71.0	433	4	US-09-740-235-32	Sequence 32, Appl
	26	14.2	71.0	575	4	US-09-669-751-108	Sequence 108, App
c	27	14.2	71.0	814	3	US-09-020-956-36	Sequence 36, Appl
c	28	14.2	71.0	814	3	US-09-030-607-36	Sequence 36, Appl
c	29	14.2	71.0	814	4	US-09-439-313-36	Sequence 36, Appl
c	30	14.2	71.0	814	4	US-09-352-616A-36	Sequence 36, Appl
c	31	14.2	71.0	814	4	US-09-232-149A-36	Sequence 36, Appl
c	32	14.2	71.0	896	4	US-09-221-017B-199	Sequence 199, App
	33	14.2	71.0	1185	4	US-09-252-991A-1931	Sequence 1931, Ap
	34	14.2	71.0	1194	2	US-08-989-478-13	Sequence 13, Appl
	35	14.2	71.0	1194	3	US-08-996-685-13	Sequence 13, Appl
	36	14.2	71.0	1428	4	US-09-328-352-1571	Sequence 1571, Ap
c	37	14.2	71.0	1449	4	US-09-107-532A-858	Sequence 858, App
	38	14.2	71.0	1597	2	US-08-989-478-9	Sequence 9, Appli
	39	14.2	71.0	1597	3	US-08-996-685-9	Sequence 9, Appli
	40	14.2	71.0	1608	2	US-08-989-478-11	Sequence 11, Appl
	41	14.2	71.0	1608	3	US-08-996-685-11	Sequence 11, Appl
	42	14.2	71.0	2011	2	US-08-989-478-6	Sequence 6, Appli
	43	14.2	71.0	2011	2	US-08-989-478-7	Sequence 7, Appli
	44	14.2	71.0	2011	3	US-08-996-685-6	Sequence 6, Appli
	45	14.2	71.0	2011	3	US-08-996-685-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-328-352-643/c
; Sequence 643, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 643
; LENGTH: 1698
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-643

Query Match 77.0%; Score 15.4; DB 4; Length 1698;
Best Local Similarity 94.1%; Pred. No. 44;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ATCGCCAAACGAGCAAC 20
||||||||||.||||||
Db 143 ATCGCCAAACCAGCAAC 127

RESULT 2
US-09-221-017B-347
; Sequence 347, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998

```
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP1182
;      FILING DATE:  31-DEC-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP1546
;      FILING DATE:  30-JAN-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PP2911
;      FILING DATE:  09-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/AU98/01023
;      FILING DATE:  10-DEC-1998
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monroy, Gladys H
;      REGISTRATION NUMBER:  32,430
;      REFERENCE/DOCKET NUMBER:  27340-20021.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-813-5600
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  347:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3032 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  UNKNOWN
;    ORIGINAL SOURCE:
;      ORGANISM:  PORYPHYROMONAS GINGIVALIS
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1...3032
US-09-221-017B-347

  Query Match             77.0%;  Score 15.4;  DB 4;  Length 3032;
  Best Local Similarity   94.1%;  Pred. No. 47;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGC 17
              ||||||||||| ||||||
Db       2855 CTGATCGCCATACGAGC 2871


RESULT 3
US-09-214-808-1/c
; Sequence 1, Application US/09214808A
; Patent No. 6475793
; GENERAL INFORMATION:
;  APPLICANT:  Rosenthal, Andre
;  APPLICANT:  Freiberg, Christoph
;  APPLICANT:  Perret, Xavier Philippe
;  APPLICANT:  Broughton, William John
;  TITLE OF INVENTION:  Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
; Patent No. 6475793
;  TITLE OF INVENTION:  Plasmid
;  FILE REFERENCE:  CARP0068
;  CURRENT APPLICATION NUMBER: US/09/214,808A
;  CURRENT FILING DATE:  1999-06-22
;  PRIOR APPLICATION NUMBER:  PCT/IB97/00950
;  PRIOR FILING DATE:  1997-07-10
;  NUMBER OF SEQ ID NOS:  1
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 536165
;   TYPE: DNA
;   ORGANISM: Rhizobium
US-09-214-808-1

  Query Match             77.0%;  Score 15.4;  DB 4;  Length 536165;
  Best Local Similarity   94.1%;  Pred. No. 80;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCA 18
              |||||||||| ||||||
Db     433213 TGATCGCCAATCGAGCA 433197


RESULT 4
US-08-311-731A-123
; Sequence 123, Application US/08311731A
; Patent No. 6583266
;  GENERAL INFORMATION:
;    APPLICANT:  SMITH, DOUGLAS
;    APPLICANT:  MAO, JEN-I
;    TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES
;    TITLE OF INVENTION:  RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
;    TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;    NUMBER OF SEQUENCES:  411
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WOLF, GREENFIELD & SACKS, P.C.
;      STREET:  600 ATLANTIC AVENUE
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,731A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GATES, EDWARD R.
;      REGISTRATION NUMBER:  31,616
;      REFERENCE/DOCKET NUMBER:  C0044/7125
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/720-3500
;      TELEFAX:  617/720-2441
;  INFORMATION FOR SEQ ID NO:  123:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36470 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  MYCOBACTERIUM LEPRAE
US-08-311-731A-123

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 36470;
  Best Local Similarity   85.0%;  Pred. No. 79;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              || ||||||| || ||||||
Db       3894 CTAATCGCCACACAAGCAAC 3913


RESULT 5
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococc
; Patent No. 6503729
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275
```

; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

```
;   NAME/KEY: misc_feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1569020)..(1569020)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1602912)..(1602912)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1603734)..(1603734)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1637998)..(1637998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1664854)..(1664855)
;   OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 1664976;
  Best Local Similarity   85.0%;  Pred. No. 97;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              || || ||||||| ||||||
Db    1565938 CTTATAGCCAAACCAGCAAC 1565919


RESULT 6
US-07-912-900-25
; Sequence 25, Application US/07912900
; Patent No. 5349125
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/912,900
;      FILING DATE:  19920713
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  543 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-07-912-900-25

  Query Match             74.0%;  Score 14.8;  DB 1;  Length 543;
  Best Local Similarity   88.9%;  Pred. No. 80;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              |||||||||||| | |||
Db         69 TGATCGCCAAACTATCAA 86


RESULT 7
US-08-285-309-25
; Sequence 25, Application US/08285309
; Patent No. 5569832
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/285,309
;      FILING DATE:  03-AUG-1994
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  543 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-285-309-25

  Query Match             74.0%;  Score 14.8;  DB 1;  Length 543;
  Best Local Similarity   88.9%;  Pred. No. 80;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              |||||||||||| | |||
Db         69 TGATCGCCAAACTATCAA 86


RESULT 8
US-08-502-046-25
; Sequence 25, Application US/08502046
; Patent No. 5861487
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Kovacic, Filippa
;    APPLICANT:  Tanaka, Yoshikazu
;    APPLICANT:  Lester, Diane R.
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING A 3,5'-
;    TITLE OF INVENTION:  HYDROXYLASE AND USES
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/502,046
;      FILING DATE:  14-JUL-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/285,309
;      FILING DATE:  03-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8633Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  543 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-502-046-25

  Query Match             74.0%;  Score 14.8;  DB 2;  Length 543;
  Best Local Similarity   88.9%;  Pred. No. 80;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              |||||||||||| | |||
Db         69 TGATCGCCAAACTATCAA 86


RESULT 9
US-08-313-075A-31
; Sequence 31, Application US/08313075A
; Patent No. 5639870
;  GENERAL INFORMATION:
;    APPLICANT:  Holton, Timothy A.
;    APPLICANT:  Cornish, Edwina C.
;    APPLICANT:  Tanaka, Yoshikazu
;    TITLE OF INVENTION:  GENETIC SEQUENCES ENCODING FLAVONOID
;    TITLE OF INVENTION:  PATHWAY ENZYMES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Scully, Scott, Murphy & Presser
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/313,075A
;      FILING DATE:  30-NOV-1994
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL 1538/92
;      FILING DATE:  27-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PL 6698/93
;      FILING DATE:  07-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PCT/AU93/00127
;      FILING DATE:  25-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9433
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  547 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..514
US-08-313-075A-31

  Query Match             74.0%;  Score 14.8;  DB 1;  Length 547;
  Best Local Similarity   88.9%;  Pred. No. 80;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              |||||||||||| | |||
Db         69 TGATCGCCAAACTATCAA 86


RESULT 10
US-09-107-532A-2028/c
; Sequence 2028, Application US/09107532A
; Patent No. 6583275
;  GENERAL INFORMATION:
;       APPLICANT: Lynn A Doucette-Stamm  and David Bush
;       TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                           ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;       NUMBER OF SEQUENCES: 7310
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;            STREET: 100 Beaver Street
;            CITY: Waltham
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02354
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: CD/ROM ISO9660
;            COMPUTER: PC
;            OPERATING SYSTEM: <Unknown>
;            SOFTWARE: ASCII
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/107,532A
;            FILING DATE: 30-Jun-1998
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/085,598
;            FILING DATE: 14 May 1998
```

; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2028:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 810 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...810
; SEQUENCE DESCRIPTION: SEQ ID NO: 2028:
US-09-107-532A-2028

Query Match 74.0%; Score 14.8; DB 4; Length 810;
Best Local Similarity 88.9%; Pred. No. 84;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGATCGCCAAACGAGCAA 19
||||| ||||||||||
Db 628 TGATCAGCAAACGAGCAA 611

RESULT 11
US-08-592-214A-17
; Sequence 17, Application US/08592214A
; Patent No. 5811536
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; TITLE OF INVENTION: Cauliflower Floral Meristem Identify
; TITLE OF INVENTION: Genes and Methods of Using Same
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,214A
; FILING DATE: 26-JAN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 1927
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4379 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2095..2098
; OTHER INFORMATION: /note= "N = one or more
; OTHER INFORMATION: nucleotides."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..4379
; OTHER INFORMATION: /note= "sequence = Arabidopsis
; OTHER INFORMATION: thaliana AP1 gene"
US-08-592-214A-17

Query Match 74.0%; Score 14.8; DB 1; Length 4379;
Best Local Similarity 88.9%; Pred. No. 1e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGATCGCCAAACGAGCAA 19
|| |||||||||| |||
Db 3961 TGGTCGCCAAACGAACAA 3978

RESULT 12
US-09-149-976-17
; Sequence 17, Application US/09149976
; Patent No. 6127123
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; TITLE OF INVENTION: Cauliflower Floral Meristem Identity
; TITLE OF INVENTION: Genes and Methods of Using Same
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/149,976
; FILING DATE: 09-SEP-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,214
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 3291
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4379 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2095..2098
; OTHER INFORMATION: /note= "N = one or more
; OTHER INFORMATION: nucleotides."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..4379
; OTHER INFORMATION: /note= "sequence = Arabidopsis
; OTHER INFORMATION: thaliana AP1 gene"
US-09-149-976-17

```
  Query Match             74.0%;  Score 14.8;  DB 3;  Length 4379;
  Best Local Similarity   88.9%;  Pred. No. 1e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              || |||||||||| |||
Db       3961 TGGTCGCCAAACGAACAA 3978


RESULT 13
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

  Query Match             74.0%;  Score 14.8;  DB 3;  Length 4403765;
  Best Local Similarity   88.9%;  Pred. No. 89;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GATCGCCAAACGAGCAAC 20
              || |||||| ||||||||
Db     682842 GACCGCCAACCGAGCAAC 682859


RESULT 14
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1

  Query Match             74.0%;  Score 14.8;  DB 3;  Length 4411529;
  Best Local Similarity   88.9%;  Pred. No. 89;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GATCGCCAAACGAGCAAC 20
              || |||||| ||||||||
Db     681399 GACCGCCAACCGAGCAAC 681416


RESULT 15
US-09-643-597-218
; Sequence 218, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 218
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-643-597-218

  Query Match             71.0%;  Score 14.2;  DB 4;  Length 405;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAA 19
              |||||  ||||| ||||||
Db         77 CTGATGTCCAAAAGAGCAA 95


Search completed: February 12, 2004, 06:08:10
Job time : 48.3171 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 04:42:03 ; Search time 113.659 Seconds
(without alignments)
648.193 Million cell updates/sec

Title: US-09-692-077D-15
Perfect score: 20
Sequence: 1 ctgatcgccaaacgagcaac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	15	US-10-001-073-15	Sequence 15, Appl
c 2	16	80.0	3325	12	US-10-369-493-44911	Sequence 44911, A
c 3	15.8	79.0	277	10	US-09-878-574-7065	Sequence 7065, Ap
4	15.8	79.0	1976	10	US-09-878-574-4681	Sequence 4681, Ap
5	15.8	79.0	5786	9	US-09-764-878-240	Sequence 240, App
6	15.8	79.0	5786	15	US-10-079-854-240	Sequence 240, App
c 7	15.4	77.0	536165	11	US-09-939-964-1	Sequence 1, Appli
c 8	15.2	76.0	705	9	US-09-770-149-235	Sequence 235, App
9	15.2	76.0	1281	15	US-10-128-714-1219	Sequence 1219, Ap
10	15.2	76.0	1284	15	US-10-128-714-2219	Sequence 2219, Ap
c 11	15.2	76.0	1359	10	US-09-938-842A-2021	Sequence 2021, Ap
c 12	15.2	76.0	1359	12	US-09-938-842A-2021	Sequence 2021, Ap
13	15.2	76.0	2031	14	US-10-108-605-206	Sequence 206, App
14	15.2	76.0	2193	9	US-09-815-242-6644	Sequence 6644, Ap
15	15.2	76.0	2241	14	US-10-108-605-116	Sequence 116, App
c 16	15.2	76.0	3184	14	US-10-108-605-324	Sequence 324, App
17	15.2	76.0	3281	15	US-10-128-714-219	Sequence 219, App
c 18	15.2	76.0	4779	9	US-09-764-878-216	Sequence 216, App
c 19	15.2	76.0	4779	15	US-10-079-854-216	Sequence 216, App
c 20	15	75.0	1449	12	US-10-369-493-32914	Sequence 32914, A
c 21	15	75.0	1449	12	US-10-369-493-33107	Sequence 33107, A
c 22	14.8	74.0	391	10	US-09-960-352-5967	Sequence 5967, Ap
c 23	14.8	74.0	391	10	US-09-983-965-958	Sequence 958, App
c 24	14.8	74.0	416	10	US-09-983-965-623	Sequence 623, App
25	14.8	74.0	987	12	US-10-369-493-28292	Sequence 28292, A
26	14.8	74.0	993	12	US-10-369-493-31049	Sequence 31049, A
c 27	14.8	74.0	1392	15	US-10-036-959B-2	Sequence 2, Appli
c 28	14.8	74.0	1527	15	US-10-156-761-350	Sequence 350, App
c 29	14.8	74.0	1741	15	US-10-142-835-29	Sequence 29, Appl
30	14.8	74.0	1795	12	US-10-369-493-36531	Sequence 36531, A
c 31	14.8	74.0	2091	12	US-10-108-260A-936	Sequence 936, App
c 32	14.8	74.0	2337	12	US-10-369-493-45576	Sequence 45576, A
33	14.8	74.0	2592	10	US-09-938-842A-1537	Sequence 1537, Ap
34	14.8	74.0	2592	12	US-09-938-842A-1537	Sequence 1537, Ap
c 35	14.8	74.0	3082	12	US-10-398-221-3557	Sequence 3557, Ap
c 36	14.8	74.0	5848	15	US-10-128-714-380	Sequence 380, App
c 37	14.8	74.0	5849	15	US-10-128-714-5380	Sequence 5380, Ap
c 38	14.8	74.0	9025608	15	US-10-156-761-1	Sequence 1, Appli
39	14.4	72.0	315	12	US-09-864-408A-3499	Sequence 3499, Ap
40	14.4	72.0	582	13	US-10-029-386-684	Sequence 684, App
41	14.4	72.0	808	13	US-10-027-632-173804	Sequence 173804,
42	14.4	72.0	808	14	US-10-027-632-173804	Sequence 173804,
c 43	14.4	72.0	1311	10	US-09-070-927A-3	Sequence 3, Appli
c 44	14.4	72.0	1455	13	US-10-138-701-60	Sequence 60, Appl
45	14.4	72.0	1501	12	US-10-398-221-1962	Sequence 1962, Ap

ALIGNMENTS

RESULT 1
US-10-001-073-15
; Sequence 15, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
; APPLICANT: Liggett, Stephen
; APPLICANT: Small, Kirsten
; TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
; FILE REFERENCE: 13073-PCT
; CURRENT APPLICATION NUMBER: US/10/001,073
; CURRENT FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-001-073-15

Query Match 100.0%; Score 20; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGATCGCCAAACGAGCAAC 20
||||||||||||||||||||
Db 1 CTGATCGCCAAACGAGCAAC 20

RESULT 2
US-10-369-493-44911/c
; Sequence 44911, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44911
; LENGTH: 3325
; TYPE: DNA
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-44911

Query Match 80.0%; Score 16; DB 12; Length 3325;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TCGCCAAACGAGCAAC 20
||||||||||||||||
Db 1189 TCGCCAAACGAGCAAC 1174

RESULT 3
US-09-878-574-7065/c
; Sequence 7065, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 7065
; LENGTH: 277
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701099380H1
US-09-878-574-7065

Query Match 79.0%; Score 15.8; DB 10; Length 277;
Best Local Similarity 89.5%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGATCGCCAAACGAGCAAC 20
|||||||||| ||| ||||
Db 277 TGATCGCCAAGCGAACAAC 259

RESULT 4
US-09-878-574-4681
; Sequence 4681, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4681
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1976)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3028-039-Q1-B2-G12
US-09-878-574-4681

Query Match 79.0%; Score 15.8; DB 10; Length 1976;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CTGATCGCCAAACGAGCAAC 20
||||||| ||| ||| ||||
Db 1174 CTGATCGACAANCGATCAAC 1193

RESULT 5
US-09-764-878-240
; Sequence 240, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 240
; LENGTH: 5786
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-878-240

Query Match 79.0%; Score 15.8; DB 9; Length 5786;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGATCGCCAAACGAGCAAC 20
| |||||||||| |||||||
Db 818 TTATCGCCAAATGAGCAAC 836

RESULT 6
US-10-079-854-240
; Sequence 240, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121C1
; CURRENT APPLICATION NUMBER: US/10/079,854
; CURRENT FILING DATE: 2002-02-22
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 240
; LENGTH: 5786
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-079-854-240

Query Match 79.0%; Score 15.8; DB 15; Length 5786;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGATCGCCAAACGAGCAAC 20
| |||||||||| |||||||
Db 818 TTATCGCCAAATGAGCAAC 836

```
RESULT 7
US-09-939-964-1/c
; Sequence 1, Application US/09939964
; Publication No. US20030054522A1
; GENERAL INFORMATION:
;  APPLICANT: Rosenthal, Andre
;  APPLICANT:  Freiberg, Christoph
;  APPLICANT:  Perret, Xavier Philippe
;  APPLICANT:  Broughton, William John
;  TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
;  TITLE OF INVENTION:  Plasmid
;  FILE REFERENCE: CARP0068
;  CURRENT APPLICATION NUMBER: US/09/939,964
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: 09/214,808
;  PRIOR FILING DATE: 1999-06-22
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 536165
;   TYPE: DNA
;   ORGANISM: Rhizobium
US-09-939-964-1

  Query Match             77.0%;  Score 15.4;  DB 11;  Length 536165;
  Best Local Similarity   94.1%;  Pred. No. 4.5e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCA 18
              |||||||||| ||||||
Db     433213 TGATCGCCAATCGAGCA 433197


RESULT 8
US-09-770-149-235/c
; Sequence 235, Application US/09770149
; Patent No. US20020059663A1
; GENERAL INFORMATION:
;  APPLICANT:  Gorlach, Jorn
;  APPLICANT:  An, Yong-Qiang
;  APPLICANT:  Hamilton, Carol M.
;  APPLICANT:  Price, Jennifer L.
;  APPLICANT:  Raines, Tracy M.
;  APPLICANT:  Yu, Yang
;  APPLICANT:  Rameaka, Joshua G.
;  APPLICANT:  Page, Amy
;  APPLICANT:  Matthew, Abraham V.
;  APPLICANT:  Ledford, Brooke L.
;  APPLICANT:  Woessner, Jeffrey P.
;  APPLICANT:  Haas, William David
;  APPLICANT:  Garcia, Carlos A.
;  APPLICANT:  Kricker, Maja
;  APPLICANT:  Slader, Ted
;  APPLICANT:  Davis, Keith R.
;  APPLICANT:  Allen, Keith
;  APPLICANT:  Hoffman, Neil
;  APPLICANT:  Hurban, Patrick
;  TITLE OF INVENTION:  Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE:  2024 (PARA-013PRV)
;  CURRENT APPLICATION NUMBER: US/09/770,149
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER:  60/178,506
;  PRIOR FILING DATE:  2000-01-27
;  NUMBER OF SEQ ID NOS:  999
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 235
;   LENGTH: 705
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(705)
;   OTHER INFORMATION: n = A,T,C or G
US-09-770-149-235

  Query Match             76.0%;  Score 15.2;  DB 9;  Length 705;
  Best Local Similarity   85.0%;  Pred. No. 3.1e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              || ||||||| ||||||| |
Db        182 CTCATCGCCAGACGAGCACC 163


RESULT 9
US-10-128-714-1219
; Sequence 1219, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1219
;   LENGTH: 1281
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-1219

  Query Match             76.0%;  Score 15.2;  DB 15;  Length 1281;
  Best Local Similarity   85.0%;  Pred. No. 3.3e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||| |||||||||| || ||
Db         37 CTGGTCGCCAAACGCGCCAC 56


RESULT 10
US-10-128-714-2219
; Sequence 2219, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
```

```
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2219
;   LENGTH: 1284
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-2219

  Query Match           76.0%;  Score 15.2;  DB 15;  Length 1284;
  Best Local Similarity   85.0%;  Pred. No. 3.3e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||| |||||||||| || ||
Db         37 CTGGTCGCCAAACGCGCCAC 56


RESULT 11
US-09-938-842A-2021/c
; Sequence 2021, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2021
;   LENGTH: 1359
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-2021

  Query Match           76.0%;  Score 15.2;  DB 10;  Length 1359;
  Best Local Similarity   85.0%;  Pred. No. 3.3e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              || ||||||| ||||||| |
Db        132 CTCATCGCCAGACGAGCACC 113


RESULT 12
US-09-938-842A-2021/c
; Sequence 2021, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
;  APPLICANT: Harper, Jeff
;  APPLICANT:  Kreps, Joel
;  APPLICANT:  Wang, Xun
;  APPLICANT:  Zhu, Tong
;  TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
;  TITLE OF INVENTION:  SAME, AND METHODS OF USE
;  FILE REFERENCE: SCRIP1300-3
;  CURRENT APPLICATION NUMBER: US/09/938,842A
;  CURRENT FILING DATE:  2001-08-24
;  PRIOR APPLICATION NUMBER: US 60/227,866
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/264,647
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/300,111
;  PRIOR FILING DATE: 2001-06-22
;  NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2021
;   LENGTH: 1359
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
US-09-938-842A-2021

  Query Match           76.0%;  Score 15.2;  DB 12;  Length 1359;
  Best Local Similarity   85.0%;  Pred. No. 3.3e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              || ||||||| ||||||| |
Db        132 CTCATCGCCAGACGAGCACC 113


RESULT 13
US-10-108-605-206
; Sequence 206, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
;  APPLICANT: Broadus, Julie
;  APPLICANT:  Stam, Lynn
;  APPLICANT:  Bachmann, Jane
;  APPLICANT:  Kamdar, Kim
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
;  TITLE OF INVENTION:  PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
;  FILE REFERENCE: 31133B
;  CURRENT APPLICATION NUMBER: US/10/108,605
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: US 09/761,142
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/176,418
;  PRIOR FILING DATE: 2000-01-14
;  NUMBER OF SEQ ID NOS: 361
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 206
;   LENGTH: 2031
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-10-108-605-206

  Query Match           76.0%;  Score 15.2;  DB 14;  Length 2031;
  Best Local Similarity   85.0%;  Pred. No. 3.4e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              | ||||||||||||  ||||
Db        695 CCGATCGCCAAACGCCCAAC 714


RESULT 14
US-09-815-242-6644
; Sequence 6644, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
```

```
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6644
;   LENGTH: 2193
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(2193)
US-09-815-242-6644

  Query Match             76.0%;  Score 15.2;  DB 9;  Length 2193;
  Best Local Similarity   85.0%;  Pred. No. 3.5e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              || |||| ||||| ||||||
Db        775 CTAATCGGCAAACAAGCAAC 794


RESULT 15
US-10-108-605-116
; Sequence 116, Application US/10108605
; Publication No. US20020160934A1
; GENERAL INFORMATION:
;  APPLICANT: Broadus, Julie
;  APPLICANT:  Stam, Lynn
;  APPLICANT:  Bachmann, Jane
;  APPLICANT:  Kamdar, Kim
;  TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
;  TITLE OF INVENTION:  PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
;  FILE REFERENCE: 31133B
;  CURRENT APPLICATION NUMBER: US/10/108,605
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: US 09/761,142
;  PRIOR FILING DATE: 2001-01-16
;  PRIOR APPLICATION NUMBER: US 60/176,418
;  PRIOR FILING DATE: 2000-01-14
;  NUMBER OF SEQ ID NOS: 361
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 116
;   LENGTH: 2241
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-10-108-605-116

  Query Match             76.0%;  Score 15.2;  DB 14;  Length 2241;
  Best Local Similarity   85.0%;  Pred. No. 3.5e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              | ||||||||||||  ||||
Db        783 CCGATCGCCAAACGCCCAAC 802
```

Search completed: February 12, 2004, 06:15:51
Job time : 120.659 secs

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 03:54:41 ; Search time 1081.71 Seconds
(without alignments)
449.373 Million cell updates/sec

Title: US-09-692-077D-15
Perfect score: 20
Sequence: 1 ctgatcgccaaacgagcaac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	17.4	87.0	275	12	BI372571	BI372571 RE59566.5
c	2	16.8	84.0	428	28	BH756337	BH756337 SALK_0535
	3	16.8	84.0	561	10	BE033222	BE033222 133577 MA
	4	16.8	84.0	700	10	BG636450	BG636450 SD14169.5
c	5	16.4	82.0	359	10	BG381314	BG381314 UI-R-CT0-
	6	16.4	82.0	477	14	N82961	N82961 TgESTzy53d1
	7	16.4	82.0	594	12	BI946431	BI946431 bEST28.5'
	8	16.4	82.0	700	28	BH962262	BH962262 odg10b08.
c	9	16.4	82.0	749	28	BH674374	BH674374 BOMNJ43TF
c	10	16.4	82.0	765	28	BH568386	BH568386 BOGYI38TF
	11	16.4	82.0	794	28	BH537924	BH537924 BOGGE84TR
c	12	16.4	82.0	806	28	BH694343	BH694343 BOHVA49TF
	13	16.4	82.0	892	29	AG071544	AG071544 Pan trogl
	14	16.4	82.0	936	14	CD255409	CD255409 AGENCOURT
	15	16.4	82.0	963	10	BG253789	BG253789 602366671
	16	16	80.0	1022	29	CNS044B1	AL273862 Tetraodon
c	17	15.8	79.0	319	29	BX293512	BX293512 Arabidops
c	18	15.8	79.0	324	29	BX293513	BX293513 Arabidops
	19	15.8	79.0	336	29	CNS00P16	AL083928 Arabidops
c	20	15.8	79.0	360	14	C65542	C65542 C65542 Yuji
	21	15.8	79.0	379	12	BJ284930	BJ284930 BJ284930
c	22	15.8	79.0	389	10	BF451274	BF451274 uz68a10.x
c	23	15.8	79.0	411	10	BF440154	BF440154 MLLG0557.
	24	15.8	79.0	417	9	AI018990	AI018990 SMEST0630
c	25	15.8	79.0	418	28	BH755131	BH755131 SALK_0484
c	26	15.8	79.0	437	10	BF072723	BF072723 NCSM4C8T7
c	27	15.8	79.0	458	29	AG256246	AG256246 Lotus jap
	28	15.8	79.0	478	12	BP026586	BP026586 BP026586
	29	15.8	79.0	481	10	BE484918	BE484918 171738 BA
c	30	15.8	79.0	487	12	BI426448	BI426448 sag03d05.
c	31	15.8	79.0	507	9	AL587025	AL587025 AL587025
c	32	15.8	79.0	513	9	AJ396379	AJ396379 AJ396379
	33	15.8	79.0	517	10	BF457549	BF457549 UI-M-BZ1-
	34	15.8	79.0	533	13	BU385646	BU385646 603860053
c	35	15.8	79.0	535	12	BI974374	BI974374 sai99h01.
	36	15.8	79.0	540	9	AI976827	AI976827 EST271509
c	37	15.8	79.0	553	12	BM525002	BM525002 sal26f07.
	38	15.8	79.0	558	12	BM328798	BM328798 PIC1_26_F
c	39	15.8	79.0	574	13	BQ770626	BQ770626 UI-M-FI0-
c	40	15.8	79.0	590	28	BH365931	BH365931 CH230-60K
	41	15.8	79.0	594	28	BH777488	BH777488 fzmb013f0
c	42	15.8	79.0	603	14	CB526603	CB526603 UI-M-FY0-
c	43	15.8	79.0	622	28	BH199984	BH199984 Sm1-46E16
	44	15.8	79.0	622	28	AQ399921	AQ399921 mgxb0003M
c	45	15.8	79.0	633	12	BJ122635	BJ122635 BJ122635

ALIGNMENTS

RESULT 1
BI372571
LOCUS BI372571 275 bp mRNA linear EST 01-AUG-2001
DEFINITION RE59566.5prime RE Drosophila melanogaster normalized Embryo pFlc-1 Drosophila melanogaster cDNA clone RE59566 5, mRNA sequence.
ACCESSION BI372571
VERSION BI372571.1 GI:15068599
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 275)
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin ,G.M.
TITLE BDGP/HHMI RE Drosophila EST Project
JOURNAL Unpublished
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798

```
          Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
          Plate: RE.595  row: F  column: 6
          High quality sequence stop: 246.
FEATURES             Location/Qualifiers
     source          1. .275
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="RE59566"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /note="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library  was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT       77 a     74 c     67 g     57 t
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 12;  Length 275;
  Best Local Similarity   94.7%;  Pred. No. 1.3e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAA 19
              |||||| ||||||||||||
Db        178 CTGATCACCAAACGAGCAA 196


RESULT 2
BH756337/c
LOCUS       BH756337                 428 bp    DNA     linear   GSS 01-MAR-2002
DEFINITION  SALK_053523.55.75.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_053523.55.75.x, genomic
            survey sequence.
ACCESSION   BH756337
VERSION     BH756337.1  GI:19038548
KEYWORDS    GSS.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 428)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .428
                     /organism="Arabidopsis thaliana"
                     /mol_type="genomic DNA"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_053523.55.75.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       95 a     86 c     101 g     146 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 28;  Length 428;
  Best Local Similarity   90.0%;  Pred. No. 3.1e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              |||||||||||  |||||||
Db        401 CTGATCGCCAATGGAGCAAC 382


RESULT 3
BE033222
LOCUS       BE033222                 561 bp    mRNA    linear   EST 09-JUL-2000
DEFINITION  133577 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION   BE033222
VERSION     BE033222.1  GI:8328231
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 561)
  AUTHORS   Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
            Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R., Quackenbush
            ,J. and Keele,J.W.
  TITLE     Porcine gene discovery by normalized cDNA-library sequencing and
            EST cluster assembly
  JOURNAL   Mamm. Genome 13 (8), 475-478 (2002)
  MEDLINE   22213789
   PUBMED   12226715
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 68  row: L  column: 5
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .561
                     /organism="Sus scrofa"
                     /mol_type="mRNA"
                     /db_xref="taxon:9823"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 1PIG"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from day 11, 13, 15, 20,
                     and 30 embryos."
BASE COUNT      130 a    179 c    135 g    117 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 10;  Length 561;
  Best Local Similarity   90.0%;  Pred. No. 3.4e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||||| || |||||||||
Db        303 CTGATCGACAGACGAGCAAC 322


RESULT 4
BG636450
```

```
LOCUS       BG636450                700 bp    mRNA    linear   EST 23-APR-2001
DEFINITION  SD14169.5prime SD Drosophila melanogaster Schneider L2 cell culture
            pOT2 Drosophila melanogaster cDNA clone SD14169 5 similar to
            CG17509: FBan0017509 located on: 2R 48E2-48E2;: 04/13/2001, mRNA
            sequence.
ACCESSION   BG636450
VERSION     BG636450.1  GI:13763987
KEYWORDS    EST.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            hit genomic AE003823: arm:2R [6997351,7260081]
            estimated-cyto:48D8-48F5: 04/13/2001
            Plate: SD.141  row: F  column: 9
            High quality sequence stop: 636.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /db_xref="taxon:7227"
                     /clone="SD14169"
                     /lab_host="DH5-alpha"
                     /clone_lib="SD Drosophila melanogaster Schneider L2 cell
                     culture pOT2"
                     /note="Vector: pOT2; Site_1: EcoRI; Site_2: Xho1; Sized
                     fractionated cDNAs were directly ligated into pOT2.
                     Plasmid cDNA library."
BASE COUNT      196 a    186 c    173 g    145 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 10;  Length 700;
  Best Local Similarity   90.0%;  Pred. No. 3.7e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCAAC 20
              ||||| |||||||||| |||
Db        232 CTGATGGCCAAACGAGAAAC 251


RESULT 5
BG381314/c
LOCUS       BG381314                359 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  UI-R-CT0-bui-g-04-0-UI.s1 UI-R-CT0 Rattus norvegicus cDNA clone
            UI-R-CT0-bui-g-04-0-UI 3', mRNA sequence.
ACCESSION   BG381314
VERSION     BG381314.1  GI:13305786
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 359)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            clones will be available through Research Genetics (www.resgen.com)
            The following repetitive elements were found in this cDNA sequence:
            73-206, >RSINE1#SINE/B4(B5)
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .359
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-CT0-bui-g-04-0-UI"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="UI-R-CT0"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CT0
                     library is a normalized library constructed from the
                     following rat brain tissues: embryonic day 17, embryonic
                     day 19, embryonic day 21, adult day 1, adult day 12, adult
                     day 75, adult day 200. For a detailed description of the
                     library from which this clone was derived, please visit
                     our web site at ratest.eng.uiowa.edu. The subtraction has
                     been previously described in (Bonaldo, Lennon and Soares,
                     Genome Research 6:791-806, 1996)
                     TAG_SEQ=None found"
BASE COUNT       73 a     96 c     71 g    119 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 10;  Length 359;
  Best Local Similarity   94.4%;  Pred. No. 4.7e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGATCGCCAAACGAGCA 18
              ||||| ||||||||||||
Db        112 CTGATTGCCAAACGAGCA 95


RESULT 6
N82961
LOCUS       N82961                  477 bp    mRNA    linear   EST 22-MAY-2000
DEFINITION  TgESTzy53d12.r1 TgRH Tachyzoite cDNA Toxoplasma gondii cDNA clone
            tgzy53d12.r1 5', mRNA sequence.
ACCESSION   N82961
VERSION     N82961.1  GI:1258714
KEYWORDS    EST.
SOURCE      Toxoplasma gondii
  ORGANISM  Toxoplasma gondii
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Toxoplasma.
REFERENCE   1  (bases 1 to 477)
  AUTHORS   Hehl,A., Manger,I., Marra,M., Sibley,L.D., Ajioka,J.A., Aslett,M.A.
            , Dietrich,N., Dubuque,T., Hillier,L., Kucaba,T., Wan,K.L.,
            Waterston,R.H. and Boothroyd,J.
  TITLE     WashU-Merck-Stanford-NIH Toxoplasma EST project
  JOURNAL   Unpublished
COMMENT     Contact: Marra M
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: toxo@watson.wustl.edu
```

```
            David Sibley at toxoest@borcim.wustl.edu for further information
            relating to organism, clone or library availability.
            Seq primer: T3
            High quality sequence stop: 301.
FEATURES             Location/Qualifiers
     source          1. .477
                     /organism="Toxoplasma gondii"
                     /mol_type="mRNA"
                     /strain="RH"
                     /db_xref="taxon:5811"
                     /clone="tgzy53d12.r1"
                     /lab_host="XL1-Blue MRF'"
                     /clone_lib="TgRH Tachyzoite cDNA"
                     /note="Vector: Lambda ZAP; Site_1: EcoRI; Site_2: XhoI;
                     Toxoplasma RH strain tachyzoites were grown in human
                     foreskin fibroblast cultures in vitro. The library was
                     constructed by K.L. Wan, Cambridge University.  cDNAs were
                     synthesized from polyA RNAs by oligo d(T) priming and
                     directionally cloned into the EcoRI to XhoI sites of the
                     Lambda ZapII vector using the ZAP-cDNA synthesis kit
                     (Stratagene). WARNING: the library contains a small
                     percentage of cDNAs derived from the human host cells."
BASE COUNT      104 a    107 c    146 g    109 t     11 others
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 14;  Length 477;
  Best Local Similarity   94.4%;  Pred. No. 5.2e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              |||| ||||||||||||||
Db        383 TGATTGCCAAACGAGCAA 400


RESULT 7
BI946431
LOCUS       BI946431                 594 bp    mRNA    linear   EST 19-OCT-2001
DEFINITION  bEST28.5' Honeybee brain cDNA library Apis mellifera cDNA 5', mRNA
            sequence.
ACCESSION   BI946431
VERSION     BI946431.1  GI:16285097
KEYWORDS    EST.
SOURCE      Apis mellifera (honeybee)
  ORGANISM  Apis mellifera
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
            Apidae; Apis.
REFERENCE   1  (bases 1 to 594)
  AUTHORS   Kucharski,R. and Maleszka,R.
  TITLE     Evaluation of differential gene expression during behavioral
            development in the honeybee using microarrays and northern blots
  JOURNAL   Genome Biol. 3 (2), research0007.1-9 (2002)
  MEDLINE   21853287
   PUBMED   11864369
COMMENT     Other_ESTs: bEST28.3'
            Contact: Maleszka R
            Visual Sciences Group
            Research School of Biological Sciences, Australian National
            University
            Biology Pl., Canberra ACT 0200, Australia
            Tel: +61 2 6125 0451
            Fax: +61 2 6125 3784
            Email: maleszka@rsbs.anu.edu.au
            Seq primer: M13 reverse.
FEATURES             Location/Qualifiers
     source          1. .594
                     /organism="Apis mellifera"
                     /mol_type="mRNA"
                     /db_xref="taxon:7460"
                     /sex="female"
                     /tissue_type="central brain, mushroom bodies, central body
                     , deuterocerebrum"
                     /dev_stage="adult worker"
                     /lab_host="E.coli SOLR"
                     /clone_lib="Honeybee brain cDNA library"
                     /note="Organ: brain; Vector: pBluescript II SK-; Site_1:
                     EcoRI; Site_2: XhoI; cDNA-harboring phagemids were
                     obtained by performing mass excision procedure on honeybee
                     brain cDNA lambda UniZap XR library (Ebert P.R. et al.,
                     1998, Insect Mol. Biol. 7(2):151-62)."
BASE COUNT      161 a    132 c    184 g    117 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 12;  Length 594;
  Best Local Similarity   94.4%;  Pred. No. 5.7e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              ||| ||||||||||||||
Db        394 TGAACGCCAAACGAGCAA 411


RESULT 8
BH962262
LOCUS       BH962262                 700 bp    DNA     linear   GSS 01-OCT-2002
DEFINITION  odg10b08.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
            sequence.
ACCESSION   BH962262
VERSION     BH962262.1  GI:23443488
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 700)
  AUTHORS   Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T., Nash
            ,W., Rabinowicz,P.D. and Wilson,R.K.
  TITLE     Whole genome shotgun reads from Brassica oleracea
  JOURNAL   Unpublished
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Plate: odg10  row: b  column: 08
            Seq primer: -21UPpOT forward
            Class: shotgun
            High quality sequence start: 20
            High quality sequence stop: 541.
FEATURES             Location/Qualifiers
     source          1. .700
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:3712"
                     /clone_lib="B.oleracea002"
                     /note="Vector: pOTw13; Whole genome shotgun library from
                     flowering buds. DNA was purified  from a crude nuclear
                     prep using Brassica oleracea TO1000DH3 buds  provided by
                     Thosmas Osborn at the University of Wisconsin. Genomic
                     DNA was provided by Pablo Rabinowicz (CSHL) and the
                     shotgun library  prepared at Washington University Genome
                     Sequencing Center."
BASE COUNT      190 a    197 c    175 g    137 t      1 others
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 28;  Length 700;
  Best Local Similarity   94.4%;  Pred. No. 6e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              || ||||||||||||||
Db        222 TGTTCGCCAAACGAGCAA 239


RESULT 9
BH674374/c
```

```
LOCUS       BH674374                749 bp    DNA     linear   GSS 19-FEB-2002
DEFINITION  BOMNJ43TF BO_2_3_KB Brassica oleracea genomic clone BOMNJ43,
            genomic survey sequence.
ACCESSION   BH674374
VERSION     BH674374.1  GI:18743024
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished
COMMENT     Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOMNJ43"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      154 a    190 c    191 g    214 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 28;  Length 749;
  Best Local Similarity   94.4%;  Pred. No. 6.2e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              || |||||||||||||||
Db        665 TGTTCGCCAAACGAGCAA 648


RESULT 10
BH568386/c
LOCUS       BH568386                765 bp    DNA     linear   GSS 14-DEC-2001
DEFINITION  BOGYI38TF BOGY Brassica oleracea genomic clone BOGYI38, genomic
            survey sequence.
ACCESSION   BH568386
VERSION     BH568386.1  GI:17820226
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 765)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished
COMMENT     Other_GSSs: BOGYI38TR
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .765
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGYI38"
                     /clone_lib="BOGY"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      158 a    161 c    225 g    221 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 28;  Length 765;
  Best Local Similarity   94.4%;  Pred. No. 6.2e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              || |||||||||||||||
Db         68 TGTTCGCCAAACGAGCAA 51


RESULT 11
BH537924
LOCUS       BH537924                794 bp    DNA     linear   GSS 14-DEC-2001
DEFINITION  BOGGE84TR BOGG Brassica oleracea genomic clone BOGGE84, genomic
            survey sequence.
ACCESSION   BH537924
VERSION     BH537924.1  GI:17776032
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished
COMMENT     Other_GSSs: BOGGE84TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .794
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGGE84"
                     /clone_lib="BOGG"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      241 a    221 c    170 g    162 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 28;  Length 794;
  Best Local Similarity   94.4%;  Pred. No. 6.3e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              || |||||||||||||||
Db        631 TGTTCGCCAAACGAGCAA 648


RESULT 12
BH694343/c
```

```
LOCUS       BH694343                806 bp    DNA     linear   GSS 19-FEB-2002
DEFINITION  BOHVA49TF BO_2_3_KB Brassica oleracea genomic clone BOHVA49,
            genomic survey sequence.
ACCESSION   BH694343
VERSION     BH694343.1  GI:18764923
KEYWORDS    GSS.
SOURCE      Brassica oleracea
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 806)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished
COMMENT     Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TF
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .806
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOHVA49"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      168 a    167 c    239 g    232 t
ORIGIN

  Query Match           82.0%;  Score 16.4;  DB 28;  Length 806;
  Best Local Similarity   94.4%;  Pred. No. 6.4e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 TGATCGCCAAACGAGCAA 19
              || ||||||||||||||||
Db        115 TGTTCGCCAAACGAGCAA 98


RESULT 13
AG071544
LOCUS       AG071544                892 bp    DNA     linear   GSS 03-NOV-2001
DEFINITION  Pan troglodytes DNA, clone: PTB-062M08.R, genomic survey sequence.
ACCESSION   AG071544
VERSION     AG071544.1  GI:16623346
KEYWORDS    GSS.
SOURCE      Pan troglodytes (chimpanzee)
  ORGANISM  Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     BAC end sequences of Library PTB
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 892)
  AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the R&D process and may have higher chance of
            clone tracking errors.
            PRIMERS
                Sequencing: M13Rev
            LIBRARY
                Vector    : pKS145
                R.Site 1  : SacI
                R.Site 2  : SacI.
FEATURES             Location/Qualifiers
     source          1. .892
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
                     /clone="PTB-062M08.R"
                     /sex="male"
                     /cell_type="lymphoblast"
                     /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT      300 a    351 c    211 g     23 t      7 others
ORIGIN

  Query Match           82.0%;  Score 16.4;  DB 29;  Length 892;
  Best Local Similarity   94.4%;  Pred. No. 6.6e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GATCGCCAAACGAGCAAC 20
              || ||||||||||||||||
Db        711 GAACGCCAAACGAGCAAC 728


RESULT 14
CD255409
LOCUS       CD255409                936 bp    mRNA    linear   EST 22-MAY-2003
DEFINITION  AGENCOURT_14164603 NICHD_XGC_Brn1 Xenopus laevis cDNA clone
            IMAGE:6951820 5', mRNA sequence.
ACCESSION   CD255409
VERSION     CD255409.1  GI:31015875
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 936)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement:
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14571  row: i  column: 03
            High quality sequence stop: 534.
FEATURES             Location/Qualifiers
     source          1. .936
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:6951820"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_XGC_Brn1"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.5 kb. Constructed by Life
                     Technologies. Note: This is a Xenopus Gene Collection (XGC
                     ) library."
```

```
BASE COUNT      270 a    190 c    225 g    251 t
ORIGIN

  Query Match            82.0%;  Score 16.4;  DB 14;  Length 936;
  Best Local Similarity   94.4%;  Pred. No. 6.7e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GATCGCCAAACGAGCAAC 20
              |||||||||||||||| |
Db        519 GATCGCCAAACGAGCAGC 536


RESULT 15
BG253789
LOCUS       BG253789                963 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602366671F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4474961 5',
            mRNA sequence.
ACCESSION   BG253789
VERSION     BG253789.1  GI:12763605
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 963)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10299  row: n  column: 18
            High quality sequence stop: 666.
FEATURES             Location/Qualifiers
     source          1. .963
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4474961"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_91"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      328 a    226 c    208 g    200 t      1 others
ORIGIN

  Query Match            82.0%;  Score 16.4;  DB 10;  Length 963;
  Best Local Similarity   94.4%;  Pred. No. 6.8e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GATCGCCAAACGAGCAAC 20
              |||||||||||||| ||||
Db        776 GATCGCCAAACGAACAAC 793


Search completed: February 12, 2004, 06:05:58
Job time : 1087.71 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 01:54:29 ; Search time 585.854 Seconds
(without alignments)
1396.582 Million cell updates/sec

Title: US-09-692-077D-16
Perfect score: 20
Sequence: 1 aaaaacgccaatgaccacag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	6	AX350504	AX350504 Sequence
2	c	20	100.0	885	9	HUMA2C2	M38742 Human alpha
3	c	20	100.0	1344	6	AX350490	AX350490 Sequence
4	c	20	100.0	1344	9	AF316895	AF316895 Homo sapi
5	c	20	100.0	1353	6	AX350489	AX350489 Sequence
6	c	20	100.0	2072	6	AR270618	AR270618 Sequence
7	c	20	100.0	2072	9	HUMADRA2RA	M34041 Human alpha
8	c	20	100.0	3274	6	AX548756	AX548756 Sequence
9	c	20	100.0	9842	9	AF005900	AF005900 Homo sapi
10		20	100.0	22842	9	AC092603	AC092603 Homo sapi
11	c	18.4	92.0	1168	4	PVI251176	AJ251176 Phoca vit
12	c	18.4	92.0	1174	4	SSC251177	AJ251177 Sus scrof
13	c	18.4	92.0	1185	10	ASP427259	AJ427259 Anomaluru
14		18.4	92.0	149169	5	AL840630	AL840630 Zebrafish
15	c	18.4	92.0	171406	2	BX510991	BX510991 Danio rer
16	c	17.4	87.0	238	11	BX323232	BX323232 Arabidops
17	c	17.4	87.0	1171	4	DDAAR2B	Y15947 Dugong dugo
18	c	17.4	87.0	1185	4	AF337541	AF337541 Tonatia b
19	c	17.4	87.0	1203	4	TTE315939	AJ315939 Tapirus t
20		17.4	87.0	45740	2	AC012972	AC012972 Drosophil
21	c	17.4	87.0	62054	9	AF161800	AF161800 Homo sapi
22		17.4	87.0	77521	8	F7H2	AC034256 Sequence
23		17.4	87.0	95359	9	AC078903	AC078903 Homo sapi
24	c	17.4	87.0	109219	9	AF159227	AF159227 Homo sapi
25	c	17.4	87.0	132063	9	AC017096	AC017096 Homo sapi
26		17.4	87.0	155132	2	AC080148	AC080148 Homo sapi
27	c	17.4	87.0	159032	9	AC083843	AC083843 Homo sapi
28	c	17.4	87.0	164752	2	AC116696	AC116696 Mus muscu
29	c	17.4	87.0	168583	3	AC104510	AC104510 Drosophil
30	c	17.4	87.0	180919	2	AC007521	AC007521 Drosophil
31	c	17.4	87.0	189652	2	BX537119	BX537119 Danio rer
32	c	17.4	87.0	283075	3	AE003485	AE003485 Drosophil
33		17.4	87.0	334028	2	AC116537	AC116537 Drosophil
34		17	85.0	193800	2	AC126994	AC126994 Rattus no
35		17	85.0	255782	2	AC106301	AC106301 Rattus no
36	c	16.8	84.0	1149	4	RAM315937	AJ315937 Rousettus
37	c	16.8	84.0	1152	10	DSA427263	AJ427263 Dipus sag
38	c	16.8	84.0	1153	4	EMA2ABREC	Y12525 Elephas max
39	c	16.8	84.0	1159	4	AHA2ABREC	Y12526 Amblysomus
40	c	16.8	84.0	1162	4	FCA251174	AJ251174 Felis cat
41	c	16.8	84.0	1172	10	ARU427256	AJ427256 Aplodonti
42	c	16.8	84.0	1176	10	TSW427267	AJ427267 Thryonomy
43	c	16.8	84.0	1180	10	CPO271336	AJ271336 Cavia por
44	c	16.8	84.0	1183	4	DRO419811	AJ419811 Desmodus
45	c	16.8	84.0	1197	4	LCR427254	AJ427254 Lepus cra

ALIGNMENTS

RESULT 1
AX350504
LOCUS AX350504 20 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 16 from Patent WO0179561.
ACCESSION AX350504
VERSION AX350504.1 GI:18616100
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liggett,S.B. and Small,K.M.
TITLE Alpha-2 adrenergic receptor polymorphisms
JOURNAL Patent: WO 0179561-A 16 25-OCT-2001;

```
          Liggett, Stephen B. (US) ; Small, Kersten M. (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT       10 a      6 c      3 g      1 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 6.2;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db          1 AAAAACGCCAATGACCACAG 20


RESULT 2
HUMA2C2/c
LOCUS       HUMA2C2                   885 bp    DNA     linear   PRI 06-MAR-1995
DEFINITION  Human alpha-2 adrenergic receptor (ADRA2C) gene, partial cds.
ACCESSION   M38742
VERSION     M38742.1  GI:177867
KEYWORDS    alpha-2 adrenergic receptor.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 885)
  AUTHORS   Chang,A.C., Ho,T.F. and Chang,N.C.
  TITLE     In vitro amplification by polymerase chain reaction of a partial
            gene encoding the third subtype of alpha-2 adrenergic receptor in
            humans
  JOURNAL   Biochem. Biophys. Res. Commun. 172 (2), 817-823 (1990)
  MEDLINE   91054503
   PUBMED   2173582
COMMENT     Original source text: Human adult neuroblastoma DNA, clone PCRA2.
            Draft entry and computer-readable sequence for [Unpublished (1990)]
            kindly submitted
            by A.C.Chang, 20-SEP-1990.
               Institute of Neuroscience
               155, Sect II, Li-Noon St.
               Taipei, Taiwan, 11221 ROC.
FEATURES             Location/Qualifiers
     source          1. .885
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="4p16.3-p15"
                     /clone="PCRA2"
                     /cell_line="neuroblastoma SK-N-SH"
                     /cell_type="neuroblastoma"
                     /tissue_type="neural"
                     /dev_stage="adult"
     gene            1. .885
                     /gene="ADRA2C"
     CDS             1. .885
                     /partial
                     /gene="ADRA2C"
                     /codon_start=1
                     /product="alpha-2-adrenergic receptor"
                     /protein_id="AAA62823.1"
                     /db_xref="GI:177868"
                     /db_xref="GDB:G00-120-540"
                     /translation="FCTSSIVHLCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWL
                     IAAVISLPPLIYKGDQGPQPRGRPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIY
                     LIAKRSNRRGPRAKGGPGQGESKQPRPDHGGALASAKLPALASVASAREVNGHSKSTG
                     EKEEGETPEDTGTRALPPSWAALPNSGQGQKEGVCGASPEDEAEEEEEEEEEEECEP
                     QAVPVSPASACSPPLQQPQGSRVLATLRGQVLLGRGVGAIGGQWWRRRAQLTREKRFT
                     FVLAVVIGVFVLCWFPFFF"
BASE COUNT      152 a    285 c    285 g    163 t
ORIGIN
```

```
  Query Match             100.0%;  Score 20;  DB 9;  Length 885;
  Best Local Similarity   100.0%;  Pred. No. 9.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db        858 AAAAACGCCAATGACCACAG 839


RESULT 3
AX350490/c
LOCUS       AX350490                1344 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 2 from Patent WO0179561.
ACCESSION   AX350490
VERSION     AX350490.1  GI:18616092
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liggett,S.B. and Small,K.M.
  TITLE     Alpha-2 adrenergic receptor polymorphisms
  JOURNAL   Patent: WO 0179561-A 2 25-OCT-2001;
            Liggett, Stephen B. (US) ; Small, Kersten M. (US)
FEATURES             Location/Qualifiers
     source          1. .1344
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      220 a    458 c    400 g    266 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 9.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1131 AAAAACGCCAATGACCACAG 1112


RESULT 4
AF316895/c
LOCUS       AF316895                1344 bp    DNA     linear   PRI 13-FEB-2001
DEFINITION  Homo sapiens alpha 2B adrenergic receptor (ADRA2B) gene, complete
            cds.
ACCESSION   AF316895
VERSION     AF316895.1  GI:12698669
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1344)
  AUTHORS   Small,K.M., Brown,K.M., Forbes,S.L. and Liggett,S.B.
  TITLE     Polymorphic deletion of three intracellular acidic residues of the
            alpha 2B-adrenergic receptor decreases G protein-coupled receptor
            kinase-mediated phosphorylation and desensitization
  JOURNAL   J. Biol. Chem. 276 (7), 4917-4922 (2001)
  MEDLINE   21265012
   PUBMED   11056163
REFERENCE   2  (bases 1 to 1344)
  AUTHORS   Small,K.M., Brown,K.M., Forbes,S.L. and Liggett,S.B.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2000) Internal Medicine, University of
            Cincinnati, 231 Bethesda Ave, Cincinnati, OH 45267, USA
FEATURES             Location/Qualifiers
     source          1. .1344
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
```

```
                /chromosome="2"
     gene       <1. .>1344
                /gene="ADRA2B"
     mRNA       <1. .>1344
                /gene="ADRA2B"
                /product="alpha 2B adrenergic receptor"
     CDS        1. .1344
                /gene="ADRA2B"
                /note="G-protein coupled receptor"
                /codon_start=1
                /product="alpha 2B adrenergic receptor"
                /protein_id="AAK01635.1"
                /db_xref="GI:12698670"
                /translation="MDHQDPYSVQATAAIAAAITFLILFTIFGNALVILAVLTSRSLR
                APQNLFLVSLAAADILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVH
                LCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPRG
                RPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIYLIAKRSNRRGPRAKGGPGQGES
                KQPRPDHGGALASAKLPALASVASAREVNGHSKSTGEKEEGETPEDTGTRALPPSWAA
                LPNSGQGQKEGVCGASPEDEAEEEEEEEECEPQAVPVSPASACSPPLQQPQGSRVLA
                TLRGQVLLGRGVGAIGGQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGA
                ICPKHCKVPHGLFQFFFWIGYCNSSLNPVIYTIFNQDFRRAFRRILCRPWTQTAW"
     variation  891^892
                /gene="ADRA2B"
                /note="compared to wild type sequence presented in GenBank
                Accession Number AF005900; polymorphic sequence lacks
                three glu residues at this location"
                /frequency="Caucasians 0.31; African-Americans 0.12"
                /replace="gaagaggag"
BASE COUNT      220 a    458 c    400 g    266 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 9;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 9.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1131 AAAAACGCCAATGACCACAG 1112


RESULT 5
AX350489/c
LOCUS       AX350489               1353 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 1 from Patent WO0179561.
ACCESSION   AX350489
VERSION     AX350489.1  GI:18616091
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Liggett,S.B. and Small,K.M.
  TITLE     Alpha-2 adrenergic receptor polymorphisms
  JOURNAL   Patent: WO 0179561-A 1 25-OCT-2001;
            Liggett, Stephen B. (US) ; Small, Kersten M. (US)
FEATURES             Location/Qualifiers
     source          1. .1353
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT      224 a    458 c    405 g    266 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 9.5;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121


RESULT 6
AR270618/c
LOCUS       AR270618               2072 bp    DNA     linear   PAT 10-APR-2003
DEFINITION  Sequence 1181 from patent US 6500938.
ACCESSION   AR270618
VERSION     AR270618.1  GI:29701852
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2072)
  AUTHORS   Au-Young,J. and Seilhamer,J.J.
  TITLE     Composition for the detection of signaling pathway gene expression
  JOURNAL   Patent: US 6500938-A 1181 31-DEC-2002;
FEATURES             Location/Qualifiers
     source          1. .2072
                     /organism="unknown"
BASE COUNT      316 a    705 c    660 g    391 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 2072;
  Best Local Similarity   100.0%;  Pred. No. 9.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1552 AAAAACGCCAATGACCACAG 1533


RESULT 7
HUMADRA2RA/c
LOCUS       HUMADRA2RA             2072 bp    DNA     linear   PRI 30-OCT-1994
DEFINITION  Human alpha-2-adrenergic receptor (alpha-2 c2) gene, complete cds.
ACCESSION   M34041
VERSION     M34041.1  GI:178197
KEYWORDS    alpha-2-adrenergic receptor; plasma membrane protein;
            receptor-coupled G protein.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2072)
  AUTHORS   Lomasney,J.W., Lorenz,W., Allen,L.F., King,K., Regan,J.W.,
            Yang-Feng,T.L., Caron,M.G. and Lefkowitz,R.J.
  TITLE     Expansion of the alpha 2-adrenergic receptor family: cloning and
            characterization of a human alpha 2-adrenergic receptor subtype,
            the gene for which is located on chromosome 2
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 87 (13), 5094-5098 (1990)
  MEDLINE   90311349
   PUBMED   2164221
COMMENT     Original source text: Human placenta DNA, clone alpha-2 C2.
            Draft entry and computer-readable sequence for [1] kindly submitted
            by J.W.Lomasney, 03-MAY-1990, for release after publication.
FEATURES             Location/Qualifiers
     source          1. .2072
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="2"
     gene            413. .1765
                     /gene="ADRA2B"
     CDS             413. .1765
                     /gene="ADRA2B"
                     /note="alpha-2-adrenergic receptor (alpha-2 C2)  old gene
                     name 'ADRA2RL1'"
                     /codon_start=1
                     /protein_id="AAA51666.1"
                     /db_xref="GI:178198"
                     /db_xref="GDB:G00-120-539"
                     /translation="MDHQDPYSVQATAAIAAAITFLILFTIFGNALVILAVLTSRSLR
                     APQNLFLVSLAAADILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVH
                     LCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPRG
                     RPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIYLIAKRSNRRGPRAKGGPGQGES
```

KQPRPDHGGALASAKLPALASVASAREVNGHSKSTGEKEEGETPEDTGTRALPPSWAA
LPNSGQGQKEGVCGASPEDEAEEEEEEEEEEECEPQAVPVSPASACSPPLQQPQGSR
VLATLRGQVLLGRGVGAIGGQWWRRRAHVTREKRFTFVLAVVIGVFVLCWFPFFFSYS
LGAICPKHCKVPHGLFQFFFWIGYCNSSLNPVIYTIFNQDFRRAFRRILCRPWTQTAW
"

BASE COUNT 316 a 705 c 660 g 391 t
ORIGIN Chromosome 2.

Query Match 100.0%; Score 20; DB 9; Length 2072;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
||||||||||||||||||||
Db 1552 AAAAACGCCAATGACCACAG 1533

RESULT 8
AX548756/c
LOCUS AX548756 3274 bp DNA linear PAT 26-NOV-2002
DEFINITION Sequence 41 from Patent WO02061087.
ACCESSION AX548756
VERSION AX548756.1 GI:25813686
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Burmer,G.C., Roush,C.L. and Brown,J.P.
TITLE Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides
JOURNAL Patent: WO 02061087-A 41 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES Location/Qualifiers
source 1. .3274
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 587 a 979 c 967 g 741 t
ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 3274;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
||||||||||||||||||||
Db 1140 AAAAACGCCAATGACCACAG 1121

RESULT 9
AF005900/c
LOCUS AF005900 9842 bp DNA linear PRI 08-JUL-1997
DEFINITION Homo sapiens alpha2B-adrenergic receptor (alpha2C2AR) gene, complete cds.
ACCESSION AF005900
VERSION AF005900.1 GI:2245627
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 9842)
AUTHORS Cayla,C., Schaak,S., Bouloumie,A., Devedjian,J.C. and Paris,H.
TITLE Alpha2C2-adrenergic receptor gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 9842)
AUTHORS Cayla,C., Schaak,S., Bouloumie,A., Devedjian,J.C. and Paris,H.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-1997) INSERM Unit 317, Institut Louis Bugnard, CHU Rangueil, Toulouse 31403, France

FEATURES Location/Qualifiers
source 1. .9842
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/clone_lib="brain fetal genomic library"
gene 5398. .6750
/gene="alpha2C2AR"
/note="alpha2C2 adrenergic receptor gene"
CDS 5398. .6750
/gene="alpha2C2AR"
/note="adrenergic receptor"
/codon_start=1
/product="alpha2B-adrenergic receptor"
/protein_id="AAB62558.1"
/db_xref="GI:2245628"
/translation="MDHQDPYSVQATAAIAAAITFLILFTIFGNALVILAVLTSRSLR
APQNLFLVSLAAADILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVH
LCAISLDRYWAVSRALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPRG
RPQCKLNQEAWYILASSIGSFFAPCLIMILVYLRIYLIAKRSNRRGPRAKGGPGQGES
KQPRPDHGGALASAKLPALASVASAREVNGHSKSTGEKEEGETPEDTGTRALPPSWAA
LPNSGQGQKEGVCGASPEDEAEEEEEEEEEEECEPQAVPVSPASACSPPLQQPQGSR
VLATLRGQVLLGRGVGAIGGQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYS
LGAICPKHCKVPHGLFQFFFWIGYCNSSLNPVIYTIFNQDFRRAFRRILCRPWTQTAW
"
polyA_signal 8649. .8654
BASE COUNT 1870 a 2747 c 3225 g 1998 t 2 others
ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 9842;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
||||||||||||||||||||
Db 6537 AAAAACGCCAATGACCACAG 6518

RESULT 10
AC092603
LOCUS AC092603 22842 bp DNA linear PRI 01-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-139J6 from 2, complete sequence.
ACCESSION AC092603 AC073396
VERSION AC092603.2 GI:16303539
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 22842)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 22842)
AUTHORS Martinka,S., Abbott,A., Hawkins,M., Elliott,G. and Doebber,A.
TITLE The sequence of Homo sapiens BAC clone RP11-139J6
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 22842)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 22842)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 22842)

```
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-OCT-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 22842)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Oct 20, 2001 this sequence version replaced gi:14916188.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0139J06
                Drafting Center: WIBR
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-401C13, 2000 bp overlap;
            the clone sequenced to the right is RP11-574O17, 2000 bp overlap.
            Actual end of this clone is at base position 48999 of RP11-574O17.

            Polymorphisms have been identified between AC013272 and AC092603.

            The sequence of AC073396 has been incorporated into AC092603.
FEATURES             Location/Qualifiers
     source          1. .22842
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-139J6"
                     /clone_lib="RPCI-11"
     repeat_region   352. .465
                     /rpt_family="MIR"
     repeat_region   1377. .1424
                     /rpt_family="(TG)n"
     repeat_region   1424. .1449
                     /rpt_family="(GA)n"
     repeat_region   1495. .1665
                     /rpt_family="Alu"
     repeat_region   1669. .1966
                     /rpt_family="Alu"
     repeat_region   2812. .2918
                     /rpt_family="MIR"
     repeat_region   3183. .3448
                     /rpt_family="CT-rich"
     repeat_region   5541. .5566
                     /rpt_family="AT_rich"
     repeat_region   5576. .5870
                     /rpt_family="Alu"
     repeat_region   5873. .6183
                     /rpt_family="Alu"
     repeat_region   6195. .6301
                     /rpt_family="MIR"
     repeat_region   6627. .6733
                     /rpt_family="L2"
     repeat_region   7278. .7415
                     /rpt_family="MIR"
     repeat_region   7698. .7723
                     /rpt_family="(GGCTG)n"
     repeat_region   8140. .8218
                     /rpt_family="G-rich"
     repeat_region   8863. .8922
                     /rpt_family="MIR"
     repeat_region   9191. .9346
                     /rpt_family="MIR"
     repeat_region   10172. .10296
                     /rpt_family="MIR"
     repeat_region   10436. .10550
                     /rpt_family="MIR"
     repeat_region   11016. .11169
                     /rpt_family="MIR"
     repeat_region   11617. .11882
                     /rpt_family="L2"
     misc_feature    12062. .12532
                     /note="match to EST AI733093 (NID:g5054206) oj40h05.x5"
     misc_feature    12062. .12257
                     /note="match to EST AA887330 (NID:g3002438) oj40h05.s1"
     misc_feature    12378. .12902
                     /note="match to EST AI791589 (NID:g5339305) oj40h05.y5"
     misc_feature    13059
                     /note="match to EST BF475329 (NID:g11546156)"
     misc_feature    13072. .13459
                     /note="match to EST BF475329 (NID:g11546156)"
     misc_feature    14030. .14244
                     /note="similar to Homo sapiens EST T39448 (NID:g647179)
                     ya06a09.r2"
     misc_feature    14060. .14247
                     /note="similar to Homo sapiens EST AL549866
                     (NID:g12886265)"
     misc_feature    14060. .14247
                     /note="similar to Homo sapiens EST AL573897
                     (NID:g12933576)"
     repeat_region   14413. .14439
                     /rpt_family="(TCC)n"
     misc_feature    14762. .14822
                     /note="similar to Homo sapiens EST AL544609
                     (NID:g12877089)"
     misc_feature    14849. .16413
                     /note="CpG_island (%GC=74.2, o/e=0.90, #CpGs=191)"
     misc_feature    14900. .15218
                     /note="similar to Homo sapiens EST AL544609
                     (NID:g12877089)"
     misc_feature    15084. .15487
                     /note="similar to Sus scrofa EST AW785035 (NID:g7841811)"
     misc_feature    15110. .15648
                     /note="similar to Rattus norvegicus EST AW916723
                     (NID:g8082456)"
     repeat_region   15747. .15832
                     /rpt_family="G-rich"
     repeat_region   18481. .18596
```

```
                     /rpt_family="L2"
     repeat_region   19266. .19675
                     /rpt_family="L2"
     repeat_region   20027. .20223
                     /rpt_family="MER1_type"
     repeat_region   20316. .20416
                     /rpt_family="L2"
     repeat_region   20537. .20973
                     /rpt_family="L2"
     misc_feature    22057. .22471
                     /note="match to EST AA830781 (NID:g2903880) oc56b01.s1"
     misc_feature    22057. .22208
                     /note="match to EST AA836522 (NID:g2910841) od22d08.s1"
BASE COUNT     4797 a   7173 c   6248 g   4624 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 9;  Length 22842;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db      14200 AAAAACGCCAATGACCACAG 14219


RESULT 11
PVI251176/c
LOCUS       PVI251176               1168 bp    DNA     linear   MAM 01-JUN-2001
DEFINITION  Phoca vitulina partial aar2B gene for alpha adrenergic receptor 2B.
ACCESSION   AJ251176
VERSION     AJ251176.1  GI:11322419
KEYWORDS    aar2B gene; alpha adrenergic receptor 2B.
SOURCE      Phoca vitulina (harbor seal)
  ORGANISM  Phoca vitulina
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Phoca.
REFERENCE   1
  AUTHORS   Madsen,O., Scally,M., Douady,C.J., Kao,D.J., DeBry,R.W., Adkins,R.,
            Amrine,H.M., Stanhope,M.J., de Jong,W.W. and Springer,M.S.
  TITLE     Parallel adaptive radiations in two major clades of placental
            mammals
  JOURNAL   Nature 409 (6820), 610-614 (2001)
  MEDLINE   21082081
   PUBMED   11214318
REFERENCE   2  (bases 1 to 1168)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1999) Madsen O., Department of Biochemistry,
            University of Nijmegen, P.O. box 9101, 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1168
                     /organism="Phoca vitulina"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9720"
     gene            1. .1168
                     /gene="aar2B"
     CDS             <1. .>1168
                     /gene="aar2B"
                     /codon_start=2
                     /product="alpha adrenergic receptor 2B"
                     /protein_id="CAC16696.1"
                     /db_xref="GI:11322420"
                     /db_xref="SPTREMBL:Q9GL07"
                     /translation="AIAAVITFLILFTIFGNALVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDQGPQPAGAPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPGAKGGPGPGESKQPRSVPTGTSTK
                     LPTLASLAASEEANGHSKPTGEKEGDTPEDPGTPALPPSWSALPSSGQSQKEGVCGAS
                     PEEEAEEEEEEEECEPQALPASPASACSPPLPQPQGSRVLATLRGQVLLGRGVGTAS
                     GQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      189 a    398 c    351 g    230 t
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 4;  Length 1168;
  Best Local Similarity   95.0%;  Pred. No. 72;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||||||||||||||||| |
Db       1084 AAAAACGCCAATGACCACGG 1065


RESULT 12
SSC251177/c
LOCUS       SSC251177               1174 bp    DNA     linear   MAM 01-JUN-2001
DEFINITION  Sus scrofa partial aar2B gene for alpha adrenergic receptor 2B.
ACCESSION   AJ251177
VERSION     AJ251177.1  GI:11322421
KEYWORDS    aar2B gene; alpha adrenergic receptor 2B.
SOURCE      Sus scrofa (pig)
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1
  AUTHORS   Madsen,O., Scally,M., Douady,C.J., Kao,D.J., DeBry,R.W., Adkins,R.,
            Amrine,H.M., Stanhope,M.J., de Jong,W.W. and Springer,M.S.
  TITLE     Parallel adaptive radiations in two major clades of placental
            mammals
  JOURNAL   Nature 409 (6820), 610-614 (2001)
  MEDLINE   21082081
   PUBMED   11214318
REFERENCE   2  (bases 1 to 1174)
  AUTHORS   Madsen,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1999) Madsen O., Department of Biochemistry,
            University of Nijmegen, P.O. box 9101, 6500 HB Nijmegen,
            NETHERLANDS
FEATURES             Location/Qualifiers
     source          1. .1174
                     /organism="Sus scrofa"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9823"
     gene            1. .1174
                     /gene="aar2B"
     CDS             <1. .>1174
                     /gene="aar2B"
                     /codon_start=2
                     /product="alpha adrenergic receptor 2B"
                     /protein_id="CAC16697.1"
                     /db_xref="GI:11322422"
                     /db_xref="SPTREMBL:Q9GL06"
                     /translation="AIAAVITFLILFTIFGNSLVILAVLTSRSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFRRTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKCIILTVWLIAAVISLPPLIYKGDPGPQPRGRPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYLIAKRSHRRGPRAKGGPGKGESKQPRPVPGEVSAS
                     AKLPTLVSQLATAGETNGCSQHTGEKDEGETSEDTGTPALPPSWPALPSSGQGQKEGV
                     CGTSPEEEAEEEEEEECEPQTLPASPASACSPPLQQPQGSRVLATLRGQVLLGRGMGT
                     SSGQWWRRRAQLTREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPHGLF"
BASE COUNT      196 a    376 c    346 g    256 t
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 4;  Length 1174;
  Best Local Similarity   95.0%;  Pred. No. 72;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||| ||||||||||||||||
Db       1090 AAACACGCCAATGACCACAG 1071


RESULT 13
ASP427259/c
LOCUS       ASP427259               1185 bp    DNA     linear   ROD 25-OCT-2002
DEFINITION  Anomalurus sp. partial A2AB gene for alpha 2B adrenergic receptor,
            exon 1.
```

```
ACCESSION   AJ427259
VERSION     AJ427259.1  GI:21655549
KEYWORDS    A2AB gene; alpha 2B adrenergic receptor.
SOURCE      Anomalurus sp.
  ORGANISM  Anomalurus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Anomaluridae;
            Anomalurus.
REFERENCE   1
  AUTHORS   Huchon,D., Madsen,O., Sibbald,M.J., Ament,K., Stanhope,M.J.,
            Catzeflis,F., de Jong,W.W. and Douzery,E.J.
  TITLE     Rodent phylogeny and a timescale for the evolution of Glires:
            evidence from an extensive taxon sampling using three nuclear genes
  JOURNAL   Mol. Biol. Evol. 19 (7), 1053-1065 (2002)
  MEDLINE   22077630
   PUBMED   12082125
REFERENCE   2  (bases 1 to 1185)
  AUTHORS   Douzery,E.J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JAN-2002) Douzery E.J.P., Institut des Sciences de
            l'Evolution, Lab. Paleonto., Paleobio., Phylogenie, Universite
            Montpellier II, Pl. E. Bataillon, 34 095 Montpellier Cedex 5,
            FRANCE
FEATURES             Location/Qualifiers
     source          1. .1185
                     /organism="Anomalurus sp."
                     /mol_type="genomic DNA"
                     /db_xref="taxon:101665"
     gene            1. .1185
                     /gene="A2AB"
     CDS             <1. .>1185
                     /gene="A2AB"
                     /codon_start=1
                     /product="alpha 2B adrenergic receptor"
                     /protein_id="CAD20297.1"
                     /db_xref="GI:21655550"
                     /translation="AIAAVTTFLILFTIFGNALVILAVLTSHSLRAPQNLFLVSLAAA
                     DILVATLIIPFSLANELLGYWYFWSTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVS
                     RALEYNSKRTPRRIKRIILIVWLIAAVISLPPLIYKGDQGPQPRGRPQCKLNQEAWYI
                     LASSIGSFFAPCLIMILVYLRIYVIAKRSNRRGPRTNGRPGQGESKQFPPVPGGAPAS
                     AKVPTLSSPLSSAGEANGHPKPPGEKEDGETPEDPGARALPSSWAALPNSGQGQKKGA
                     CGASAEEEAEEEEEEEEEEECEPQAVPVSPASVCSQPLQQPQGSRVLATLRGQVLLGR
                     HVGTAGGQWWRRRAQLSREKRFTFVLAVVIGVFVLCWFPFFFSYSLGAICPQHCKVPH
                     GLF"
     exon            <1. .>1185
                     /gene="A2AB"
                     /number=1
BASE COUNT      202 a    391 c    344 g    247 t      1 others
ORIGIN
```

```
  Query Match             92.0%;  Score 18.4;  DB 10;  Length 1185;
  Best Local Similarity   95.0%;  Pred. No. 72;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||| ||||||||||||||||
Db       1101 AAAGACGCCAATGACCACAG 1082
```

```
RESULT 14
AL840630
LOCUS       AL840630              149169 bp    DNA     linear   VRT 24-MAY-2003
DEFINITION  Zebrafish DNA sequence from clone CH211-279L11 in linkage group 17,
            complete sequence.
ACCESSION   AL840630
VERSION     AL840630.15  GI:31071445
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 149169)
  AUTHORS   Lloyd,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 23, 2003 this sequence version replaced gi:30722485.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Repeat names
            beginning 'Dr' were identified by the Recon repeat discovery system
            (Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr'
            were identified by Rick Waterman (Stephen Johnson lab, WashU).  For
            further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-279L11 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .149169
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-279L11"
                     /clone_lib="CHORI-211"
BASE COUNT    47889 a  26273 c  26492 g  48515 t
ORIGIN
```

```
  Query Match             92.0%;  Score 18.4;  DB 5;  Length 149169;
  Best Local Similarity   95.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||||| ||||||||||||||
Db      21021 AAAAACACCAATGACCACAG 21040
```

```
RESULT 15
BX510991/c
LOCUS       BX510991              171406 bp    DNA     linear   HTG 22-MAY-2003
DEFINITION  Danio rerio clone RP71-44C4, *** SEQUENCING IN PROGRESS ***, 8
            unordered pieces.
ACCESSION   BX510991
VERSION     BX510991.3  GI:31071389
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 171406)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (21-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On May 23, 2003 this sequence version replaced gi:30962299.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: bZ44C4
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 167404 bases at least Q40
            Consensus quality: 168324 bases at least Q30
            Consensus quality: 169018 bases at least Q20
            Insert size: 170706; sum-of-contigs
            Insert size: 169513; 3.5% error; agarose-fp
            Quality coverage: 7.49x in Q20 bases; sum-of-contigs Quality
            coverage: 7.55x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 8 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *      1    56569: contig of 56569 bp in length
            *  56570    56669: gap of 100 bp
            *  56670    62067: contig of 5398 bp in length
            *  62068    62167: gap of 100 bp
            *  62168    88720: contig of 26553 bp in length
            *  88721    88820: gap of 100 bp
            *  88821   134708: contig of 45888 bp in length
            * 134709   134808: gap of 100 bp
            * 134809   142476: contig of 7668 bp in length
            * 142477   142576: gap of 100 bp
            * 142577   154106: contig of 11530 bp in length
            * 154107   154206: gap of 100 bp
            * 154207   164114: contig of 9908 bp in length
            * 164115   164214: gap of 100 bp
            * 164215   171406: contig of 7192 bp in length.
FEATURES             Location/Qualifiers
     source          1. .171406
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="RP71-44C4"
                     /clone_lib="RPCI-71"
     misc_feature    1. .56569
                     /note="assembly_fragment:01943
                     fragment_chain:1
                     clone_end:T7
                     vector_side:left"
     misc_feature    56670. .62067
                     /note="assembly_fragment:00525
                     fragment_chain:1"
     misc_feature    62168. .88720
                     /note="assembly_fragment:02082
                     fragment_chain:1"
     misc_feature    88821. .134708
                     /note="assembly_fragment:01880
                     fragment_chain:1"
     misc_feature    134809. .142476
                     /note="assembly_fragment:01032
                     fragment_chain:1"
     misc_feature    142577. .154106
                     /note="assembly_fragment:01420
                     fragment_chain:1"
     misc_feature    154207. .164114
                     /note="assembly_fragment:00067
                     fragment_chain:1"
     misc_feature    164215. .171406
                     /note="assembly_fragment:00689
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:right"
BASE COUNT    57035 a  29367 c  29252 g  55051 t    701 others
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 2;  Length 171406;
  Best Local Similarity   95.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||||| ||||||||||||||
Db      33134 AAAAACACCAATGACCACAG 33115


Search completed: February 12, 2004, 04:41:51
Job time : 589.854 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 01:53:19 ; Search time 105.122 Seconds
(without alignments)
513.582 Million cell updates/sec

Title: US-09-692-077D-16
Perfect score: 20
Sequence: 1 aaaaacgccaatgaccacag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	23	AAI99910	Human alpha-2BAR g
c	2	20	100.0	1344	22	AAD04761	Human alpha2B-adre
c	3	20	100.0	1344	23	AAI99906	Human alpha-2BAR t
c	4	20	100.0	1344	24	AAD44388	Human alpha-2B-adr
c	5	20	100.0	1353	22	AAD04762	Human alpha2B-adre
c	6	20	100.0	1353	23	AAI99905	Human alpha-2BAR t
c	7	20	100.0	1353	24	AAD44389	Human alpha-2B-adr
c	8	20	100.0	2064	12	AAQ14151	Human alpha 2 beta
c	9	20	100.0	2064	18	AAT59499	Human alpha-2b adr
c	10	20	100.0	2072	25	ACA56583	Human signalling p
c	11	20	100.0	3274	25	ABZ42624	Human alpha 2b-adr
c	12	17.4	87.0	965	24	ABN98529	Arabidopsis thalia
	13	17.4	87.0	4910	23	ABL03674	Drosophila melanog
c	14	16.8	84.0	1441	18	AAT48444	Human TGF-beta-lik
c	15	16.8	84.0	2177	18	AAT48442	Human TGF-beta-lik
c	16	16.8	84.0	2178	18	AAT48443	Human TGF-beta-lik
c	17	16.8	84.0	8902	24	ABL51551	Human NAG-1 promot
	18	16.4	82.0	788	24	ABQ68453	Listeria monocytog
c	19	16.4	82.0	996	22	AAH65708	C glutamicum codin
	20	16.4	82.0	4504	24	ABQ70977	Listeria monocytog
	21	16.4	82.0	349980	22	AAH68526	C glutamicum codin
	22	16.4	82.0	2944528	24	ABA03041	Listeria monocytog
c	23	16.4	82.0	2944528	24	ABA03041	Listeria monocytog
	24	15.8	79.0	192	23	ABV31952	Human prostate exp
	25	15.8	79.0	192	23	ABV40901	Human prostate exp
c	26	15.8	79.0	201	23	ABV01487	Human prostate exp
	27	15.8	79.0	201	23	ABV01490	Human prostate exp
	28	15.8	79.0	201	23	ABV01543	Human prostate exp
c	29	15.8	79.0	201	23	ABV01635	Human prostate exp
	30	15.8	79.0	201	23	ABV01725	Human prostate exp
	31	15.8	79.0	201	23	ABV01746	Human prostate exp
	32	15.8	79.0	201	23	ABV04990	Human prostate exp
c	33	15.8	79.0	201	23	ABV10656	Human prostate exp
	34	15.8	79.0	201	23	ABV10659	Human prostate exp
c	35	15.8	79.0	201	23	ABV10681	Human prostate exp
	36	15.8	79.0	201	23	ABV10712	Human prostate exp
c	37	15.8	79.0	201	23	ABV10804	Human prostate exp
	38	15.8	79.0	201	23	ABV10828	Human prostate exp
	39	15.8	79.0	201	23	ABV10837	Human prostate exp
c	40	15.8	79.0	201	23	ABV10857	Human prostate exp
	41	15.8	79.0	201	23	ABV10894	Human prostate exp
	42	15.8	79.0	201	23	ABV10915	Human prostate exp
	43	15.8	79.0	201	23	ABV12301	Human prostate exp
	44	15.8	79.0	201	23	ABV14159	Human prostate exp
c	45	15.8	79.0	203	23	ABV01688	Human prostate exp

ALIGNMENTS

RESULT 1
AAI99910
ID AAI99910 standard; DNA; 20 BP.
XX
AC AAI99910;
XX
DT 18-FEB-2002 (first entry)
XX
DE Human alpha-2BAR genotyping PCR primer SEQ ID NO 16.
XX
KW Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW polymorphic site; allelic variant; cardiovascular disease;
KW central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW phosphorylation; inositol phosphate; alpha-2BAR; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200179561-A2.
XX
PD 25-OCT-2001.
XX
PF 17-APR-2001; 2001WO-US12575.
XX
PR 17-APR-2000; 2000US-0551744.
PR 10-AUG-2000; 2000US-0636259.
PR 19-OCT-2000; 2000US-0692077.
XX
PA (LIGG/) LIGGETT S B.
PA (SMAL/) SMALL K M.
XX
PI Liggett SB, Small KM;

```
XX
DR   WPI; 2001-611728/70.
XX
PT   Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT   determining whether an individual is at increased risk of developing a
PT   disease associated with the corresponding receptor comprises detecting
PT   a polymorphic site -
XX
PS   Claim 10; Page 112; 163pp; English.
XX
CC   The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC   receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC   (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC   alpha2A or alpha2C or fragment or complement of; and
CC   (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC   of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC   or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC   positions 961-972 of (III). The method may be used for genotyping an
CC   alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC   whether an individual is at increased risk of developing a disease
CC   associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC   polymorphic site which correlate to disease selected from cardiovascular
CC   disease, central nervous system disease and combinations of these. In
CC   addition, the technique may be used to predict an individual's response
CC   to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC   norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC   combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC   rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC   these) by detecting the polymorphic site and correlating the site to a
CC   predetermined response (where the response is correlated to adenylyl
CC   cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC   levels). The present sequence is that of a human alpha-2BAR PCR primer,
CC   useful for the genotyping methods of the invention.
XX
SQ   Sequence 20 BP; 10 A; 6 C; 3 G; 1 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 23;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db          1 AAAAACGCCAATGACCACAG 20


RESULT 2
AAD04761/c
ID   AAD04761 standard; DNA; 1344 BP.
XX
AC   AAD04761;
XX
DT   04-JUL-2001  (first entry)
XX
DE   Human alpha2B-adrenoceptor (alpha2B-AR) variant gene.
XX
KW   Human; cardiant; gene therapy; alpha2B-adrenoceptor; alpha2B-AR;
KW   glutamic acid repeat; intracellular loop; chromosome 2; catecholamine;
KW   norepinephrine; epinephrine; therapy; vascular contraction; variant;
KW   coronary artery; coronary heat disease; CHD; chronic angina pectoris;
KW   acute myocardial infarction; AMI; Prinzmetal's variant; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1344
FT                   /*tag=  a
FT                   /product= "Human alpha2B-adrenoceptor (alpha2B-AR)
FT                   variant protein"
XX
PN   WO200129082-A1.
XX
PD   26-APR-2001.
XX
PF   20-OCT-2000; 2000WO-FI00913.
XX
PR   22-OCT-1999;   99US-0422985.
XX
PA   (JUVA-) JUVANTIA PHARMA LTD OY.
XX
PI   Snapir A,  Heinonen P,  Alhopuro P,  Karvonen M,  Koulu M,  Pesonen U;
PI   Scheinin M,  Salonen JT,  Tuomainen T,  Lakka TA,  Nyyssoenen K;
PI   Salonen R,  Kauhanen J,  Valkonen V;
XX
DR   WPI; 2001-300318/31.
DR   P-PSDB; AAE00989.
XX
PT   New DNA molecule encoding variant specific adrenoceptor protein with
PT   deletion of specific amino acids located in the third intracellular
PT   loop of the polypeptide, for treating vascular contraction of coronary
PT   arteries -
XX
PS   Claim 3; Page 24-26; 37pp; English.
XX
CC   The present sequence is a gene encoding human alpha2B-adrenoceptor
CC   (alpha2B-AR) variant protein. Alpha2B-AR has a glutamic acid repeat
CC   element (amino acids 298-309) of 12 glutamates, in an acidic stretch of
CC   18 amino acids (amino acids 294-311), located in the third intracellular
CC   loop of the receptor polypeptide. The variant is obtained by deletion of
CC   three glutamates from the Glu repeat (amino acids 307-309). Alpha2B-AR
CC   gene is located on chromosome 2. Alpha2-AR mediate many of the
CC   physiological effects of the catecholamines, norepinephrine and
CC   epinephrine. An antagonist of alpha2B-adrenoceptor is useful for treating
CC   a mammal suffering from vascular contraction of coronary arteries and a
CC   disease involving vascular contraction of coronary arteries which is
CC   clinically expressed as coronary heat disease (CHD), unstable chronic
CC   angina pectoris which is clinically expressed as Prinzmetal's variant
CC   form or acute myocardial infarction (AMI). Alpha2B-AR gene is used in
CC   gene therapy.
XX
SQ   Sequence 1344 BP; 219 A; 459 C; 400 G; 266 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 22;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1131 AAAAACGCCAATGACCACAG 1112


RESULT 3
AAI99906/c
ID   AAI99906 standard; DNA; 1344 BP.
XX
AC   AAI99906;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Human alpha-2BAR third intracellular loop variant encoding DNA.
XX
KW   Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW   polymorphic site; allelic variant; cardiovascular disease;
KW   central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW   phosphorylation; inositol phosphate; alpha-2BAR; chromosome 2; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1344
FT                   /*tag=  a
FT                   /product= "alpha-2BAR"
FT                   /note= "sequence is deleted for a 9 nucleotide
FT                          polymorphic site found at nucleotides 901-909
FT                          of the wildtype alpha-2BAR protein (AAI99905)"
XX
PN   WO200179561-A2.
```

```
XX
PD   25-OCT-2001.
XX
PF   17-APR-2001; 2001WO-US12575.
XX
PR   17-APR-2000; 2000US-0551744.
PR   10-AUG-2000; 2000US-0636259.
PR   19-OCT-2000; 2000US-0692077.
XX
PA   (LIGG/) LIGGETT S B.
PA   (SMAL/) SMALL K M.
XX
PI   Liggett SB,  Small KM;
XX
DR   WPI; 2001-611728/70.
DR   P-PSDB; AAM52118.
XX
PT   Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT   determining whether an individual is at increased risk of developing a
PT   disease associated with the corresponding receptor comprises detecting
PT   a polymorphic site -
XX
PS   Claim 5; Page 144-145; 163pp; English.
XX
CC   The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC   receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC   (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC   alpha2A or alpha2C or fragment or complement of; and
CC   (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC   of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC   or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC   positions 961-972 of (III). The method may be used for genotyping an
CC   alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC   whether an individual is at increased risk of developing a disease
CC   associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC   polymorphic site which correlate to disease selected from cardiovascular
CC   disease, central nervous system disease and combinations of these. In
CC   addition, the technique may be used to predict an individual's response
CC   to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC   norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC   combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC   rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC   these) by detecting the polymorphic site and correlating the site to a
CC   predetermined response (where the response is correlated to adenylyl
CC   cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC   levels). The present sequence is that of the third intracellular loop of
CC   the human alpha-2BAR variant, the sequence is deleted for a 9 nucleotide
CC   polymorphic site found at nucleotides 901-909 of the wildtype gene
CC   (AAI99905).
XX
SQ   Sequence 1344 BP; 220 A; 458 C; 400 G; 266 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 23;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1131 AAAAACGCCAATGACCACAG 1112


RESULT 4
AAD44388/c
ID   AAD44388 standard; DNA; 1344 BP.
XX
AC   AAD44388;
XX
DT   13-DEC-2002  (first entry)
XX
DE   Human alpha-2B-adrenoceptor variant DNA.
XX
KW   Human; hypertension; alpha-2B-adrenoceptor; AR; antihypertensive;
KW   hypertension; hypotensive; variant; gene; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1344
FT                   /*tag=  a
FT                   /product= "Human alpha-2B-adrenoceptor variant
FT                   protein"
XX
PN   WO200266617-A1.
XX
PD   29-AUG-2002.
XX
PF   13-FEB-2002; 2002WO-FI00113.
XX
PR   20-FEB-2001; 2001FI-0000323.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Salonen J;
XX
DR   WPI; 2002-667063/71.
DR   P-PSDB; AAE26633.
XX
PT   Detecting a risk of hypertension and targeting treatment in a subject
PT   by determining the pattern of alleles encoding a variant
PT   alpha-2-adrenoceptor  -
XX
PS   Disclosure; Page 24-26; 35pp; English.
XX
CC   The invention relates to a method for detecting a risk of hypertension
CC   by determining the pattern of alleles encoding a variant alpha-2B-
CC   adrenoceptor (AR) protein. The methods and compositions of the invention
CC   are useful for detecting risks and targeting treatment for hypertension.
CC   The kit is also useful for selecting for clinical drug trials testing
CC   the antihypertensive effect of compounds. The present sequence is human
CC   alpha-2B-adrenoceptor variant DNA.
XX
SQ   Sequence 1344 BP; 219 A; 459 C; 400 G; 266 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 24;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1131 AAAAACGCCAATGACCACAG 1112


RESULT 5
AAD04762/c
ID   AAD04762 standard; DNA; 1353 BP.
XX
AC   AAD04762;
XX
DT   04-JUL-2001  (first entry)
XX
DE   Human alpha2B-adrenoceptor (alpha2B-AR) gene.
XX
KW   Human; cardiant; gene therapy; alpha2B-adrenoceptor; alpha2B-AR;
KW   glutamic acid repeat; intracellular loop; chromosome 2; catecholamine;
KW   norepinephrine; epinephrine; therapy; vascular contraction;
KW   coronary artery; coronary heat disease; CHD; chronic angina pectoris;
KW   acute myocardial infarction; AMI; Prinzmetal's variant; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1353
FT                   /*tag=  a
FT                   /product= "Human alpha2B-adrenoceptor (alpha2B-AR)
FT                   protein"
```

```
XX
PN   WO200129082-A1.
XX
PD   26-APR-2001.
XX
PF   20-OCT-2000; 2000WO-FI00913.
XX
PR   22-OCT-1999;   99US-0422985.
XX
PA   (JUVA-) JUVANTIA PHARMA LTD OY.
XX
PI   Snapir A,  Heinonen P,  Alhopuro P,  Karvonen M,  Koulu M,  Pesonen U;
PI   Scheinin M,  Salonen JT,  Tuomainen T,  Lakka TA,  Nyyssoenen K;
PI   Salonen R,  Kauhanen J,  Valkonen V;
XX
DR   WPI; 2001-300318/31.
DR   P-PSDB; AAE00990.
XX
PT   New DNA molecule encoding variant specific adrenoceptor protein with
PT   deletion of specific amino acids located in the third intracellular
PT   loop of the polypeptide, for treating vascular contraction of coronary
PT   arteries -
XX
PS   Disclosure; Page 27-29; 37pp; English.
XX
CC   The present sequence is a gene encoding human alpha2B-adrenoceptor
CC   (alpha2B-AR) protein. Alpha2B-AR has a glutamic acid repeat element
CC   (amino acids 298-309) of 12 glutamates, in an acidic stretch of 18 amino
CC   acids (amino acids 294-311), located in the third intracellular loop of
CC   the receptor polypeptide. Alpha2B-AR gene is located on chromosome 2.
CC   Alpha2-AR mediate many of the physiological effects of the
CC   catecholamines, norepinephrine and epinephrine. An antagonist of
CC   alpha2B-adrenoceptor is useful for treating a mammal suffering from
CC   vascular contraction of coronary arteries and a disease involving
CC   vascular contraction of coronary arteries which is clinically expressed
CC   as coronary heat disease (CHD), unstable chronic angina pectoris which is
CC   clinically expressed as Prinzmetal's variant form or acute myocardial
CC   infarction (AMI). Alpha2B-AR gene is used in gene therapy.
XX
SQ   Sequence 1353 BP; 223 A; 459 C; 405 G; 266 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 22;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121


RESULT 6
AAI99905/c
ID   AAI99905 standard; DNA; 1353 BP.
XX
AC   AAI99905;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Human alpha-2BAR third intracellular loop encoding DNA.
XX
KW   Human; genotyping; alpha-2B; alpha-2A; alpha-2C; adrenergic receptor;
KW   polymorphic site; allelic variant; cardiovascular disease;
KW   central nervous system disease; adenylyl cyclase; MAP kinase activity;
KW   phosphorylation; inositol phosphate; alpha-2BAR;
KW   GenBank Accession AF009500; chromosome 2; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1353
FT                   /*tag=  a
FT                   /product= "alpha-2BAR"
FT                   /note= "sequence includes a 9 nucleotide polymorphic site
FT                   at nucleotides 901-909 absent in the alpha-2BAR
FT                   variant (AAI99906)"
XX
PN   WO200179561-A2.
XX
PD   25-OCT-2001.
XX
PF   17-APR-2001; 2001WO-US12575.
XX
PR   17-APR-2000; 2000US-0551744.
PR   10-AUG-2000; 2000US-0636259.
PR   19-OCT-2000; 2000US-0692077.
XX
PA   (LIGG/) LIGGETT S B.
PA   (SMAL/) SMALL K M.
XX
PI   Liggett SB,  Small KM;
XX
DR   WPI; 2001-611728/70.
DR   P-PSDB; AAM52117.
XX
PT   Genotyping an alpha-2B, 2A, or 2C adrenergic receptor gene useful for
PT   determining whether an individual is at increased risk of developing a
PT   disease associated with the corresponding receptor comprises detecting
PT   a polymorphic site -
XX
PS   Claim 4; Page 144; 163pp; English.
XX
CC   The invention relates to genotyping an alpha-2B, 2A, or 2C adrenergic
CC   receptor gene (I)-(III) by detecting a polymorphic site, comprising;
CC   (a) obtaining a sample having a polynucleotide encoding an alpha-2B,
CC   alpha2A or alpha2C or fragment or complement of; and
CC   (b) detecting a polymorphic site comprising nucleotide positions 901-909
CC   of (I), a site comprising cytosine or guanine at position 753 of (IIV)
CC   or a site comprising (A) (ggggcggggccg) or (B) (ggggcggctgag) at
CC   positions 961-972 of (III). The method may be used for genotyping an
CC   alpha2B, alpha2A or alpha2C receptor gene and further used to determine
CC   whether an individual is at increased risk of developing a disease
CC   associated with alpha2B, alpha2A or alpha2, comprising detecting a
CC   polymorphic site which correlate to disease selected from cardiovascular
CC   disease, central nervous system disease and combinations of these. In
CC   addition, the technique may be used to predict an individual's response
CC   to an alpha2B, alpha2A, or alpha2C agonist (e.g. epinephrine,
CC   norepinephrine, clonidine, oxymetazoline, guanabenz, UK14304, BHT933 and
CC   combinations of these) or antagonist (e.g. yohimbine, prazosin, ARC 239,
CC   rauwolscine, idazoxan, tolazoline, phentolamine and combinations of
CC   these) by detecting the polymorphic site and correlating the site to a
CC   predetermined response (where the response is correlated to adenylyl
CC   cyclase, MAP kinase activity, phosphorylation or inositol phosphate
CC   levels). The present sequence is that of the third intracellular loop of
CC   the human alpha-2BAR (GenBank Accession AF009500), the sequence includes
CC   a 9 nucleotide polymorphic site at nucleotides 901-909, absent in the
CC   alpha-2BAR variant (AAI99906).
XX
SQ   Sequence 1353 BP; 224 A; 458 C; 405 G; 266 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 23;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121


RESULT 7
AAD44389/c
ID   AAD44389 standard; DNA; 1353 BP.
XX
AC   AAD44389;
XX
DT   13-DEC-2002  (first entry)
XX
```

```
DE   Human alpha-2B-adrenoceptor gene.
XX
KW   Human; hypertension; alpha-2B-adrenoceptor; AR; antihypertensive;
KW   hypertension; hypotensive; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1353
FT                   /*tag= a
FT                   /product= "Human alpha-2B-adrenoceptor protein"
XX
PN   WO200266617-A1.
XX
PD   29-AUG-2002.
XX
PF   13-FEB-2002; 2002WO-FI00113.
XX
PR   20-FEB-2001; 2001FI-0000323.
XX
PA   (JURI-) JURILAB LTD OY.
XX
PI   Salonen J;
XX
DR   WPI; 2002-667063/71.
DR   P-PSDB; AAE26634.
XX
PT   Detecting a risk of hypertension and targeting treatment in a subject
PT   by determining the pattern of alleles encoding a variant
PT   alpha-2-adrenoceptor -
XX
PS   Disclosure; Page 27-29; 35pp; English.
XX
CC   The invention relates to a method for detecting a risk of hypertension
CC   by determining the pattern of alleles encoding a variant alpha-2B-
CC   adrenoceptor (AR) protein. The methods and compositions of the invention
CC   are useful for detecting risks and targeting treatment for hypertension.
CC   The kit is also useful for selecting for clinical drug trials testing
CC   the antihypertensive effect of compounds. The present sequence is human
CC   alpha-2B-adrenoceptor gene.
XX
SQ   Sequence 1353 BP; 223 A; 459 C; 405 G; 266 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 24;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 5.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121


RESULT 8
AAQ14151/c
ID   AAQ14151 standard; DNA; 2064 BP.
XX
AC   AAQ14151;
XX
DT   06-JAN-1992  (first entry)
XX
DE   Human alpha 2 beta adrenergic receptor gene.
XX
KW   Neurotransmission; adrenaline; epinephrine; NGC-alpha2beta; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             288..1752
FT                   /*tag= a
XX
PN   US5053337-A.
XX
PD   01-OCT-1991.
XX
PF   30-OCT-1989;   89US-0428856.
XX
PR   30-OCT-1989;   89US-0428856.
XX
PA   (NEUR-) NEUROGENETIC CORP.
XX
PI   Weinshank RL,  Hartig PR;
XX
DR   WPI; 1991-310087/42.
DR   P-PSDB; AAR14149.
XX
PT   Isolated DNA encoding human adrenergic receptor - for detecting
PT   nucleic acids encoding alpha, 2-beta adrenergic receptor, for
PT   screening drugs.
XX
PS   Claim 1; Fig 2; 15pp; English.
XX
CC   Clone NGC-alpha2beta was isolated from a human spleen genomic
CC   library by screening with a fragment of the human 5-HT1A receptor
CC   gene. The gene can be used to express recombinant receptor protein
CC   which can be used to produce antibodies for inhibition of receptor
CC   function.
XX
SQ   Sequence 2064 BP; 319 A; 696 C; 659 G; 390 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 12;  Length 2064;
  Best Local Similarity   100.0%;  Pred. No. 6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1538 AAAAACGCCAATGACCACAG 1519


RESULT 9
AAT59499/c
ID   AAT59499 standard; DNA; 2064 BP.
XX
AC   AAT59499;
XX
DT   25-MAR-2003  (updated)
DT   06-MAY-1997  (first entry)
XX
DE   Human alpha-2b adrenergic receptor genomic DNA clone.
XX
KW   Alpha-2b adrenergic receptor; adrenoceptor; adrenaline;
KW   epinephrine; signal transduction; neurotransmitter; ligand; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             288..1751
FT                   /*tag= a
XX
PN   US5595880-A.
XX
PD   21-JAN-1997.
XX
PF   22-OCT-1992;   92US-0965040.
XX
PR   30-OCT-1989;   89US-0428856.
PR   30-MAY-1991;   91US-0707604.
PR   22-OCT-1992;   92US-0965040.
XX
PA   (SYNA-) SYNAPTIC PHARM CORP.
XX
PI   Hartig PR,  Weinshank RL;
XX
DR   WPI; 1997-107576/10.
DR   P-PSDB; AAW11804.
XX
PT   Assay for alpha-2b adrenergic receptor ligands - using membranes of
```

PT cells expressing recombinant receptor
XX
PS Disclosure; Fig 2A-E; 16pp; English.
XX
CC A genomic DNA clone (AAT59499) codes for human alpha-2B adrenergic
CC receptor (AAW11804), a member of the rhodopsin-like signal transducer
CC family. It was isolated from a human spleen genomic library in the
CC lambda vector Charon 28 by screening with a 1.6 kb fragment of the
CC human 5-hydroxytryptamine receptor gene. Plasmid pNGC-alpha-2B
CC comprising DNA encoding the alpha-2B adrenoceptor is deposited as
CC ATCC 68144. Vectors have been adapted to allow prodn. of alpha-2B
CC adrenoceptor in bacterial, yeast or mammalian cells; transfected
CC Ltk- cells, designated L-NGC-alpha-2B, are deposited as ATCC CRL
CC 10275. Membranes of such cells can used in novel methods to
CC identify drugs which specifically interact with, and bind to, the
CC alpha-2B adrenergic receptor.
CC (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 2064 BP; 319 A; 696 C; 659 G; 390 T; 0 other;

Query Match 100.0%; Score 20; DB 18; Length 2064;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
||||||||||||||||||||
Db 1538 AAAAACGCCAATGACCACAG 1519

RESULT 10
ACA56583/c
ID ACA56583 standard; cDNA; 2072 BP.
XX
AC ACA56583;
XX
DT 06-JUN-2003 (first entry)
XX
DE Human signalling pathway polynucleotide probe SEQ ID NO 1181.
XX
KW Human; probe; ss; array element; Parkinson's disease;
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS Homo sapiens.
XX
PN US6500938-B1.
XX
PD 31-DEC-2002.
XX
PF 30-JAN-1998; 98US-0016434.
XX
PR 30-JAN-1998; 98US-0016434.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Seilhamer JJ;
XX
DR WPI; 2003-352189/33.
XX
PT Combination of polynucleotide probes, useful as array elements in a
PT microarray for monitoring the expression of a number of target
PT polynucleotides -
XX
PS Claim 1; SEQ ID NO 1181; 65pp; English.
XX
CC The invention relates to a combination which, comprises a number of
CC polynucleotide probes comprising a sequence selected from one of the 1490
CC sequences mentioned in the specification. The combination is useful as an
CC array element in a microarray for monitoring the expression of a number
CC of target polynucleotides. The microarray is particularly useful in the
CC diagnosis and treatment of cancer and immunopathology and neuropathology.
CC The microarray is useful in diagnostics and treatment regimens, drug
CC discovery and development, toxicological and carcinogenicity studies,
CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signaling pathway populations which can be used to diagnose
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC and Parkinson's disease. The present sequence represents a polynucleotide
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.
XX
SQ Sequence 2072 BP; 316 A; 705 C; 660 G; 391 T; 0 other;

Query Match 100.0%; Score 20; DB 25; Length 2072;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
||||||||||||||||||||
Db 1552 AAAAACGCCAATGACCACAG 1533

RESULT 11
ABZ42624/c
ID ABZ42624 standard; DNA; 3274 BP.
XX
AC ABZ42624;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human alpha 2b-adrenoceptor nucleotide SEQ ID NO:41.
XX
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200261087-A2.
XX
PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US50107.
XX
PR 19-DEC-2000; 2000US-257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
DR WPI; 2003-046718/04.
DR P-PSDB; ABP81780.
XX
PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating
PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
PT cancer or autoimmune diseases -
XX
PS Disclosure; Fig 1; 523pp; English.
XX
CC The present invention describes antigenic peptides (I) comprising:

```
CC   (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC   acids. Also described: (1) an assay for the detection of a particular
CC   G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC   and (2) an isolated antibody having high specificity and high affinity
CC   or avidity for a particular GPCR. (I) can be used as GPCR modulators and
CC   in gene therapy. The antigenic peptides for GPCRs are useful in detecting
CC   an antibody against a particular GPCR, and in the production of specific
CC   antibodies. The peptides and antibodies are also useful for detecting the
CC   presence or absence of corresponding GPCRs. The antigenic peptides for
CC   GPCRs and antibodies are useful for diagnosing and designing drugs for
CC   treating immune-related diseases, growth-related diseases, cell
CC   regeneration-related disease, immunological-related cell proliferative
CC   diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC   atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC   osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC   inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC   disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC   anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC   loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC   hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC   any other disorder in which GPCRs are involved. The antibodies may be
CC   used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC   GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 3274 BP; 587 A; 979 C; 967 G; 741 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 25;  Length 3274;
  Best Local Similarity   100.0%;  Pred. No. 6.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121


RESULT 12
ABN98529/c
ID   ABN98529 standard; DNA; 965 BP.
XX
AC   ABN98529;
XX
DT   01-AUG-2002  (first entry)
XX
DE   Arabidopsis thaliana expressed polynucleotide SEQ ID NO 297.
XX
KW   Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
KW   disease; crop; thale cress; tolerance factor; insect; pathogen;
KW   nutrition; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   US2002023281-A1.
XX
PD   21-FEB-2002.
XX
PF   26-JAN-2001; 2001US-0770445.
XX
PR   27-JAN-2000; 2000US-178472P.
XX
PA   (GORL/) GORLACH J.
PA   (ANYY/) AN Y.
PA   (HAMI/) HAMILTON C M.
PA   (PRIC/) PRICE J L.
PA   (RAIN/) RAINES T M.
PA   (YUYY/) YU Y.
PA   (RAME/) RAMEAKA J G.
PA   (PAGE/) PAGE A.
PA   (MATH/) MATHEW A V.
PA   (LEDF/) LEDFORD B L.
PA   (WOES/) WOESSNER J P.
PA   (HAAS/) HAAS W D.
PA   (GARC/) GARCIA C A.
PA   (KRIC/) KRICKER M.
PA   (SLAT/) SLATER T.
PA   (DAVI/) DAVIS K R.
PA   (ALLE/) ALLEN K.
PA   (HOFF/) HOFFMAN N.
PA   (HURB/) HURBAN P.
XX
PI   Gorlach J,  An Y,  Hamilton CM,  Price JL,  Raines TM,  Yu Y;
PI   Rameaka JG,  Page A,  Mathew AV,  Ledford BL,  Woessner JP,  Haas WD;
PI   Garcia CA,  Kricker M,  Slater T,  Davis KR,  Allen K,  Hoffman N;
PI   Hurban P;
XX
DR   WPI; 2002-400781/43.
XX
PT   New Arabidopsis thaliana nucleic acid for identifying homologous genes,
PT   producing compositions that modulate the expression or function of its
PT   encoded protein, and mapping functional regions of protein -
XX
PS   Claim 1; SEQ ID NO 297; 49pp + Sequence Listing; English.
XX
CC   The invention relates to an Arabidopsis thaliana nucleic acid (I)
CC   comprising a sequence capable of hybridising under stringent conditions
CC   to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
CC   given in the specification or its fragment. A polypeptide (II) encoded by
CC   (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
CC   genetically modified cell (IV) comprising an exogenous nucleic acid, is
CC   useful for screening a candidate agent for its biological effect. (I) is
CC   useful in identifying homologous or related genes, in producing
CC   compositions that modulate the expression or function of its encoded
CC   protein, mapping functional regions of the protein and in studying
CC   associated physiological pathways. (I) is also useful for the genetic
CC   manipulation of cells, particularly plant cells. (I) is also useful in
CC   screening assays of various plant strains to determine the strains that
CC   are best capable of withstanding a particular disease or environmental
CC   stress. (II) and (III) are useful for screening of biologically active
CC   agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
CC   pathways. The screened agents are useful in improved methods of treating
CC   crops to prevent or treat disease. (II) are also useful in screening
CC   programs to identify agents that mimic or enhance the action of tolerance
CC   factors. Such agents are useful in improved methods of treating crops to
CC   enhance their tolerance to environmental stress. (I) is also useful
CC   for enhancing or inhibiting production of a biosynthetic product in a
CC   plant. (III) is useful for identifying other mediators that may induce
CC   expression of proteins of interest, for establishing the extent to which
CC   any specific insect and/or pathogen is responsible for damage to a
CC   particular plant, for identifying other mediators that enhance or induce
CC   tolerance to environmental stress, for identifying factors involved in
CC   biosynthetic pathways of nutritional, commercial, or medicinal value and
CC   for identifying productions of nutritional, commercial or medicinal
CC   value. (IV) is useful in the study of genetic function and regulation,
CC   for alteration of the cellular metabolism and for screening compounds
CC   that may affect the biological function of the gene or gene products.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from USPTO
CC   at seqdata.uspto.gov/sequence.html?DocID=999909770445.
XX
SQ   Sequence 965 BP; 218 A; 314 C; 161 G; 268 T; 4 other;

  Query Match             87.0%;  Score 17.4;  DB 24;  Length 965;
  Best Local Similarity   94.7%;  Pred. No. 95;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACAG 20
              |||| ||||||||||||||
Db        219 AAAAGGCCAATGACCACAG 201


RESULT 13
ABL03674
ID   ABL03674 standard; cDNA; 4910 BP.
XX
AC   ABL03674;
XX
```

```
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 5504.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB59571.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 5504; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 4910 BP; 1279 A; 1083 C; 1134 G; 1414 T; 0 other;

  Query Match             87.0%;  Score 17.4;  DB 23;  Length 4910;
  Best Local Similarity   94.7%;  Pred. No. 1.1e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACA 19
              |||||||| ||||||||||
Db       2828 AAAAACGCTAATGACCACA 2846


RESULT 14
AAT48444/c
ID   AAT48444 standard; cDNA; 1441 BP.
XX
AC   AAT48444;
XX
DT   11-APR-1997  (first entry)
XX
DE   Human TGF-beta-like cytokine pCL13 variant u2.
XX
KW   pCL13; clone 13; transforming growth factor-beta; TGF-beta;
KW   immunosuppressive; cell differentiation; antiproliferative;
KW   wound healing; vulnerary; ischaemic injury; inflammation; cancer;
KW   autoimmune disease; fibrotic disease; diagnosis; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             234..1160
FT                   /*tag=  a
XX
PN   WO9700958-A1.
XX
PD   09-JAN-1997.
XX
PF   24-JUN-1996;   96WO-AU00386.
XX
PR   09-FEB-1996;   96AU-0007983.
PR   22-JUN-1995;   95AU-0003706.
PR   23-AUG-1995;   95AU-0004990.
XX
PA   (SVIN-) ST VINCENT'S HOSPITAL SYDNEY LTD.
XX
PI   Bootcov M,  Breit S;
XX
DR   WPI; 1997-087385/08.
DR   P-PSDB; AAW10668.
XX
PT   Isolated nucleic acid encoding pCL13, new TGF-beta family member -
PT   has immunosuppressant, cell differentiation promoting and
PT   anti:proliferative activities, for treatment of wounds, ischaemic
PT   injury, cancer, inflammatory disease etc.
XX
PS   Example 17; Fig 20A; 73pp; English.
XX
CC   cDNA clones b2, h1, u2, f1, a1, b1, d2 and dd2 (AAT48442-49) are
CC   variants of clone 13 (AAT48438), which codes for pCL13 (AAW10662), a
CC   novel TGF-beta-like cytokine whose activities include
CC   immunosuppression, down-regulation of leucocyte extravasation and
CC   motility, promotion of cell differentiation and inhibition of
CC   proliferation.  The clones were obtd. from a foetal lung library
CC   using a portion of the coding region of clone 13 (see also AAT48438)
CC   as probe.  They show substantial variation in the 5' untranslated
CC   region but only minor differences in the coding region.  Translated
CC   coding regions are given in AAW10666-73.  While the variants may be
CC   important in e.g. transcriptional reglation they are untranslated
CC   and hence cannot affect bioactivity.
XX
SQ   Sequence 1441 BP; 300 A; 451 C; 428 G; 262 T; 0 other;

  Query Match             84.0%;  Score 16.8;  DB 18;  Length 1441;
  Best Local Similarity   90.0%;  Pred. No. 1.9e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||| || ||||||||||||
Db        110 AAACACTCCAATGACCACAG 91


RESULT 15
AAT48442/c
ID   AAT48442 standard; cDNA; 2177 BP.
XX
AC   AAT48442;
XX
DT   11-APR-1997  (first entry)
XX
DE   Human TGF-beta-like cytokine pCL13 variant b2.
XX
KW   pCL13; clone 13; transforming growth factor-beta; TGF-beta;
KW   immunosuppressive; cell differentiation; antiproliferative;
KW   wound healing; vulnerary; ischaemic injury; inflammation; cancer;
KW   autoimmune disease; fibrotic disease; diagnosis; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             993..1919
FT                   /*tag=  a
XX
PN   WO9700958-A1.
XX
```

PD 09-JAN-1997.
XX
PF 24-JUN-1996; 96WO-AU00386.
XX
PR 09-FEB-1996; 96AU-0007983.
PR 22-JUN-1995; 95AU-0003706.
PR 23-AUG-1995; 95AU-0004990.
XX
PA (SVIN-) ST VINCENT'S HOSPITAL SYDNEY LTD.
XX
PI Bootcov M, Breit S;
XX
DR WPI; 1997-087385/08.
DR P-PSDB; AAW10666.
XX
PT Isolated nucleic acid encoding pCL13, new TGF-beta family member -
PT has immunosuppressant, cell differentiation promoting and
PT anti:proliferative activities, for treatment of wounds, ischaemic
PT injury, cancer, inflammatory disease etc.
XX
PS Example 17; Fig 20A; 73pp; English.
XX
CC cDNA clones b2, h1, u2, f1, a1, b1, d2 and dd2 (AAT48442-49) are
CC variants of clone 13 (AAT48438), which codes for pCL13 (AAW10662), a
CC novel TGF-beta-like cytokine whose activities include
CC immunosuppression, down-regulation of leucocyte extravasation and
CC motility, promotion of cell differentiation and inhibition of
CC proliferation. The clones were obtd. from a foetal lung library
CC using a portion of the coding region of clone 13 (see also AAT48438)
CC as probe. They show substantial variation in the 5' untranslated
CC region but only minor differences in the coding region. Translated
CC coding regions are given in AAW10666-73. While the variants may be
CC important in e.g. transcriptional reglation they are untranslated
CC and hence cannot affect bioactivity.
XX
SQ Sequence 2177 BP; 462 A; 645 C; 609 G; 461 T; 0 other;

Query Match 84.0%; Score 16.8; DB 18; Length 2177;
Best Local Similarity 90.0%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 AAAAACGCCAATGACCACAG 20
              ||| || ||||||||||||
Db        869 AAACACTCCAATGACCACAG 850
```

Search completed: February 12, 2004, 04:01:34
Job time : 108.122 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 04:01:41 ; Search time 27.3171 Seconds
(without alignments)
323.156 Million cell updates/sec

Title: US-09-692-077D-16
Perfect score: 20
Sequence: 1 aaaaacgccaatgaccacag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	20	100.0	2072	4	US-09-016-434-1181	Sequence 1181, Ap
	2	16	80.0	36941	4	US-08-311-731A-130	Sequence 130, App
c	3	15.2	76.0	1593	4	US-09-134-001C-1673	Sequence 1673, Ap
	4	15.2	76.0	1779	4	US-09-328-352-312	Sequence 312, App
c	5	15.2	76.0	5820	3	US-09-029-213B-7	Sequence 7, Appli
	6	15.2	76.0	6755	3	US-08-931-999-4	Sequence 4, Appli
	7	15.2	76.0	55827	4	US-09-813-133A-3	Sequence 3, Appli
c	8	14.8	74.0	215	4	US-09-177-650-103	Sequence 103, App
	9	14.8	74.0	268	3	US-09-221-298-106	Sequence 106, App
	10	14.8	74.0	291	4	US-09-313-294A-6186	Sequence 6186, Ap
c	11	14.8	74.0	24	4	US-08-169-715-1	Sequence 1, Appli
c	12	14.8	74.0	1422	4	US-09-328-352-3388	Sequence 3388, Ap
c	13	14.8	74.0	124884	4	US-09-661-596A-76	Sequence 76, Appl
	14	14.4	72.0	1788	4	US-09-328-352-1004	Sequence 1004, Ap
c	15	14.4	72.0	1858	2	US-08-359-705B-7	Sequence 7, Appli
c	16	14.4	72.0	1858	2	US-08-286-846A-7	Sequence 7, Appli
c	17	14.4	72.0	1858	2	US-08-457-880A-7	Sequence 7, Appli
c	18	14.4	72.0	1858	3	US-08-444-622A-7	Sequence 7, Appli
c	19	14.4	72.0	1858	3	US-08-942-562-7	Sequence 7, Appli
c	20	14.4	72.0	1858	3	US-09-156-923-7	Sequence 7, Appli
c	21	14.4	72.0	2715	2	US-08-359-705B-5	Sequence 5, Appli
c	22	14.4	72.0	2715	2	US-08-286-846A-5	Sequence 5, Appli
c	23	14.4	72.0	2715	2	US-08-457-880A-5	Sequence 5, Appli
c	24	14.4	72.0	2715	3	US-08-444-622A-5	Sequence 5, Appli
c	25	14.4	72.0	2715	3	US-08-942-562-5	Sequence 5, Appli
c	26	14.4	72.0	2715	3	US-09-156-923-5	Sequence 5, Appli
c	27	14.4	72.0	2940	1	US-08-286-305A-8	Sequence 8, Appli
c	28	14.4	72.0	2940	2	US-08-441-104A-8	Sequence 8, Appli
c	29	14.4	72.0	2940	2	US-08-440-816A-8	Sequence 8, Appli
c	30	14.4	72.0	2940	3	US-09-417-381A-8	Sequence 8, Appli
	31	14.2	71.0	186	4	US-09-313-294A-1872	Sequence 1872, Ap
c	32	14.2	71.0	273	4	US-09-134-001C-2061	Sequence 2061, Ap
c	33	14.2	71.0	273	4	US-09-107-532A-583	Sequence 583, App
	34	14.2	71.0	278	4	US-09-313-294A-885	Sequence 885, App
	35	14.2	71.0	294	4	US-09-134-001C-2032	Sequence 2032, Ap
c	36	14.2	71.0	390	4	US-09-134-001C-2684	Sequence 2684, Ap
	37	14.2	71.0	426	1	US-07-661-610C-12	Sequence 12, Appl
c	38	14.2	71.0	534	4	US-09-328-352-937	Sequence 937, App
	39	14.2	71.0	609	4	US-09-134-001C-2780	Sequence 2780, Ap
c	40	14.2	71.0	840	4	US-09-328-352-1188	Sequence 1188, Ap
	41	14.2	71.0	865	1	US-08-341-568-4	Sequence 4, Appli
	42	14.2	71.0	865	2	US-08-911-020-4	Sequence 4, Appli
c	43	14.2	71.0	1008	4	US-09-107-532A-3178	Sequence 3178, Ap
c	44	14.2	71.0	1031	4	US-09-397-787-157	Sequence 157, App
c	45	14.2	71.0	1341	2	US-08-945-848-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-016-434-1181/c
; Sequence 1181, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1181:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2072 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g178197
US-09-016-434-1181

Query Match 100.0%; Score 20; DB 4; Length 2072;
Best Local Similarity 100.0%; Pred. No. 0.63;

```
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1552 AAAAACGCCAATGACCACAG 1533


RESULT 2
US-08-311-731A-130
; Sequence 130, Application US/08311731A
; Patent No. 6583266
;  GENERAL INFORMATION:
;    APPLICANT:  SMITH, DOUGLAS
;    APPLICANT:  MAO, JEN-I
;    TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES
;    TITLE OF INVENTION:  RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
;    TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;    NUMBER OF SEQUENCES:  411
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  WOLF, GREENFIELD & SACKS, P.C.
;      STREET:  600 ATLANTIC AVENUE
;      CITY:  BOSTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,731A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GATES, EDWARD R.
;      REGISTRATION NUMBER:  31,616
;      REFERENCE/DOCKET NUMBER:  C0044/7125
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/720-3500
;      TELEFAX:  617/720-2441
;  INFORMATION FOR SEQ ID NO:  130:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  36941 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  MYCOBACTERIUM LEPRAE
US-08-311-731A-130

  Query Match             80.0%;  Score 16;  DB 4;  Length 36941;
  Best Local Similarity   100.0%;  Pred. No. 97;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ACGCCAATGACCACAG 20
              ||||||||||||||||
Db      26433 ACGCCAATGACCACAG 26448


RESULT 3
US-09-134-001C-1673/c
; Sequence 1673, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1673
;   LENGTH: 1593
;   TYPE: DNA
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1673

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 1593;
  Best Local Similarity   85.0%;  Pred. No. 1.4e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              || ||| |||||||| ||||
Db       1170 AATAACTCCAATGACTACAG 1151


RESULT 4
US-09-328-352-312
; Sequence 312, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 312
;   LENGTH: 1779
;   TYPE: DNA
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-312

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 1779;
  Best Local Similarity   85.0%;  Pred. No. 1.4e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||  |||||||||||| ||
Db       1377 AAAGCCGCCAATGACCATAG 1396


RESULT 5
US-09-029-213B-7/c
; Sequence 7, Application US/09029213B
; Patent No. 6180098
;  GENERAL INFORMATION:
;    APPLICANT:  CHRISTIAN, Peter D.
;    TITLE OF INVENTION:  RECOMBINANT HELICOVERPA BACULOVIRUSES
;    TITLE OF INVENTION:  EXPRESSING HETEROLOGOUS DNA
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDermott, Will & Emery
;      STREET:  600 13th Street, NW
;      CITY:  Washington
;      STATE:  District of Columbia
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/029,213B
;      FILING DATE:  31-AUG-1998
;      CLASSIFICATION:  424
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Joseph Hyosuk Kim
;      REGISTRATION NUMBER:  41,425
;      REFERENCE/DOCKET NUMBER:  50179-048
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-756-8000
;      TELEFAX:  202-756-8087
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5820 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1388..1973
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2083..2820
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  4074..4875
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  complement (5..317)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  complement (651..1194)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  complement (2821..3790)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  complement (5004..5820)
US-09-029-213B-7

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 5820;
  Best Local Similarity   85.0%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              || ||||||| ||| |||||
Db        110 AATAACGCCATTGATCACAG 91


RESULT 6
US-08-931-999-4
; Sequence 4, Application US/08931999
; Patent No. 6043219
;  GENERAL INFORMATION:
;    APPLICANT:  Iandolo, John J.
;    APPLICANT:  Crupper, Scott S.
;    TITLE OF INVENTION:  Broad Spectrum Chemotherapeutic Peptide
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hovey, Williams, Timmons & Collins
;      STREET:  2405 Grand Boulevard, Suite 400
;      CITY:  Kansas City
;      STATE:  Missouri
;      COUNTRY:  U.S.A.
;      ZIP:  64108
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/931,999
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/710,561
;      FILING DATE:  19-SEP-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Collins, John M.
;      REGISTRATION NUMBER:  26,262
;      REFERENCE/DOCKET NUMBER:  25043-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  816/474-9050
;      TELEFAX:  816/474-9057
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6755 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Staphylococcus aureus
;      STRAIN:  UT0007
US-08-931-999-4

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 6755;
  Best Local Similarity   85.0%;  Pred. No. 1.8e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||| || ||||||||
Db       2263 AAAAACGAGAAGGACCACAG 2282


RESULT 7
US-09-813-133A-3
; Sequence 3, Application US/09813133A
; Patent No. 6455294
; GENERAL INFORMATION:
;  APPLICANT: GAN, Weiniu et al
;  TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
;  TITLE OF INVENTION:  NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
;  TITLE OF INVENTION:  USES THEREOF
;  FILE REFERENCE: CL001173
;  CURRENT APPLICATION NUMBER: US/09/813,133A
;  CURRENT FILING DATE:  2001-06-06
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 55827
;   TYPE: DNA
;   ORGANISM: Human
US-09-813-133A-3

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 55827;
  Best Local Similarity   85.0%;  Pred. No. 2.5e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||  ||||| |||||||||
Db      21190 AAAGGCGCCACTGACCACAG 21209


RESULT 8
US-09-177-650-103/c
; Sequence 103, Application US/09177650
; Patent No. 6413719
; GENERAL INFORMATION:
;  APPLICANT: Leppert, Mark F.
;  APPLICANT: Singh, Nanda
;  APPLICANT: Charlier, Carole
;  TITLE OF INVENTION: KCNQ2 AND KCNQ3 - POTASSIUM CHANNEL GENES WHICH ARE
;  TITLE OF INVENTION: MUTATED IN BENIGN FAMILIAL NEONATAL CONVULSIONS (BFNC)
;  TITLE OF INVENTION: AND OTHER EPILEPSIES
;  FILE REFERENCE: 2323-134
;  CURRENT APPLICATION NUMBER: US/09/177,650
```

```
;  CURRENT FILING DATE: 1998-10-23
;  EARLIER APPLICATION NUMBER: 60/063,147
;  EARLIER FILING DATE: 1997-10-24
;  NUMBER OF SEQ ID NOS: 129
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
;   LENGTH: 215
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-177-650-103

  Query Match           74.0%;  Score 14.8;  DB 4;  Length 215;
  Best Local Similarity  88.9%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACA 19
              |||||  ||||||||||
Db        146 AAAACCACAATGACCACA 129


RESULT 9
US-09-221-298-106
; Sequence 106, Application US/09221298
; Patent No. 6284241
; GENERAL INFORMATION:
;  APPLICANT: Xu, Jiangchun
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
;  TITLE OF INVENTION: OF COLON CANCER
;  FILE REFERENCE: 210121.471
;  CURRENT APPLICATION NUMBER: US/09/221,298
;  CURRENT FILING DATE: 1998-12-23
;  NUMBER OF SEQ ID NOS: 112
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 106
;   LENGTH: 268
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (18)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (73)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (96)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (113)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (129)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (161)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (195)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (219)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (247)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (250)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (255)
;   OTHER INFORMATION: Where n is a, c, g or t
;   FEATURE:
;   NAME/KEY: modified_base
;   LOCATION: (257)
;   OTHER INFORMATION: Where n is a, c, g or t
US-09-221-298-106

  Query Match           74.0%;  Score 14.8;  DB 3;  Length 268;
  Best Local Similarity  88.9%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCAC 18
              |||||||||||||| | |
Db        200 AAAAACGCCAATGAACCC 217


RESULT 10
US-09-313-294A-6186
; Sequence 6186, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 6186
;   LENGTH: 291
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700351368H1
;   NAME/KEY: unsure
;   LOCATION: 264, 286
;   OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6186

  Query Match           74.0%;  Score 14.8;  DB 4;  Length 291;
  Best Local Similarity  88.9%;  Pred. No. 1.6e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACA 19
              ||||||  |||||||||  |
Db        172 AAAACGACAATGACCATA 189


RESULT 11
US-08-169-715-1/c
; Sequence 1, Application US/08169715
; Patent No. 6300056
;  GENERAL INFORMATION:
;    APPLICANT:  Irvine, Bruce D.
;    APPLICANT:  Horn, Thomas
;    APPLICANT:  Chang, Chu-An
;    TITLE OF INVENTION:  HIV PROBES FOR USE IN SOLUTION PHASE
;    TITLE OF INVENTION:  SANDWICH HYBRIDIZATION ASSAYS
;    NUMBER OF SEQUENCES:  63
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  545 Middlefield Road, Suite 200
;      CITY:  Menlo Park
```

```
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94025
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/169,715
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/813,583
;       FILING DATE:  18-DEC-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Monroy, Gladys H.
;       REGISTRATION NUMBER:  32,430
;       REFERENCE/DOCKET NUMBER:  22300-20150.00
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-813-5600
;       TELEFAX:  415-327-2951
;       TELEX:  706141
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  24 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-169-715-1

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 24;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+02;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 ACGCCAATGACC 16
              | ||  | ||||
Db         24 AGGCATAGGACC 13


RESULT 12
US-09-328-352-3388/c
; Sequence 3388, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3388
;   LENGTH: 1422
;   TYPE: DNA
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-3388

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 1422;
  Best Local Similarity   88.9%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AAACGCCAATGACCACAG 20
              ||||| || |||||||||
Db       1178 AAACGGCACTGACCACAG 1161


RESULT 13
US-09-661-596A-76/c
; Sequence 76, Application US/09661596A
; Patent No. 6528066
; GENERAL INFORMATION:
;  APPLICANT: Grose, Charles
;  APPLICANT: Santos, Richard
;  TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE
;  FILE REFERENCE: 140.0011 0101
;  CURRENT APPLICATION NUMBER: US/09/661,596A
;  CURRENT FILING DATE:  2000-09-14
;  PRIOR APPLICATION NUMBER: US 60/153,779
;  PRIOR FILING DATE: 1999-09-14
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 76
;   LENGTH: 124884
;   TYPE: DNA
;   ORGANISM: Varicella zoster
US-09-661-596A-76

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 124884;
  Best Local Similarity   88.9%;  Pred. No. 4.3e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AAACGCCAATGACCACAG 20
              | |||||||| ||||||
Db      46128 ATACGCCAATGTCCACAG 46111


RESULT 14
US-09-328-352-1004
; Sequence 1004, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1004
;   LENGTH: 1788
;   TYPE: DNA
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-1004

  Query Match             72.0%;  Score 14.4;  DB 4;  Length 1788;
  Best Local Similarity   93.8%;  Pred. No. 3.5e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACC 16
              ||||| ||||||||||
Db       1156 AAAAAAGCCAATGACC 1171


RESULT 15
US-08-359-705B-7/c
; Sequence 7, Application US/08359705B
; Patent No. 5844092
;  GENERAL INFORMATION:
;    APPLICANT:  Presta, Leonard G.
;    APPLICANT:  Shelton, David L.
;    APPLICANT:  Urfer, Roman
;    TITLE OF INVENTION:  Human trk Receptors and Neurotrophic Factor Inhibitors
;    NUMBER OF SEQUENCES:  41
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/359,705B
; FILING DATE: 20-Dec-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286846
; FILING DATE: 08/10/94
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/215139
; FILING DATE: 03/18/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0873P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1858 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-359-705B-7

Query Match 72.0%; Score 14.4; DB 2; Length 1858;
Best Local Similarity 93.8%; Pred. No. 3.6e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          5 ACGCCAATGACCACAG 20
              | ||||||||||||||
Db       1529 ATGCCAATGACCACAG 1514
```

Search completed: February 12, 2004, 06:08:16
Job time : 33.3171 secs

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 04:42:03 ; Search time 113.659 Seconds
(without alignments)
648.193 Million cell updates/sec

Title: US-09-692-077D-16
Perfect score: 20
Sequence: 1 aaaaacgccaatgaccacag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	15	US-10-001-073-16	Sequence 16, Appl
c 2	20	100.0	1344	9	US-09-825-923-1	Sequence 1, Appli
c 3	20	100.0	1344	15	US-10-077-870-1	Sequence 1, Appli
c 4	20	100.0	1344	15	US-10-001-073-2	Sequence 2, Appli
c 5	20	100.0	1353	9	US-09-825-923-3	Sequence 3, Appli
c 6	20	100.0	1353	15	US-10-077-870-3	Sequence 3, Appli
c 7	20	100.0	1353	15	US-10-001-073-1	Sequence 1, Appli
c 8	20	100.0	2072	12	US-10-305-720-1181	Sequence 1181, Ap
c 9	20	100.0	3274	15	US-10-225-567A-41	Sequence 41, Appl
c 10	17.4	87.0	965	9	US-09-770-445-297	Sequence 297, App
c 11	16.8	84.0	1767	12	US-10-369-493-42551	Sequence 42551, A
12	16.4	82.0	788	12	US-10-398-221-1266	Sequence 1266, Ap
c 13	16.4	82.0	996	10	US-09-738-626-743	Sequence 743, App
c 14	16.4	82.0	1282	12	US-10-292-798-1745	Sequence 1745, Ap
c 15	16.4	82.0	1282	13	US-10-017-161-2099	Sequence 2099, Ap
16	16.4	82.0	4504	12	US-10-398-221-3790	Sequence 3790, Ap
17	16.4	82.0	3309400	10	US-09-738-626-1	Sequence 1, Appli
c 18	16	80.0	766	13	US-10-027-632-156514	Sequence 156514,
c 19	16	80.0	766	14	US-10-027-632-156514	Sequence 156514,
20	16	80.0	802	13	US-10-027-632-156512	Sequence 156512,
21	16	80.0	802	13	US-10-027-632-156513	Sequence 156513,
22	16	80.0	802	14	US-10-027-632-156512	Sequence 156512,
23	16	80.0	802	14	US-10-027-632-156513	Sequence 156513,
24	15.8	79.0	492	13	US-10-027-632-83956	Sequence 83956, A
25	15.8	79.0	492	13	US-10-027-632-83957	Sequence 83957, A
26	15.8	79.0	492	13	US-10-027-632-298547	Sequence 298547,
27	15.8	79.0	492	13	US-10-027-632-298548	Sequence 298548,
28	15.8	79.0	492	14	US-10-027-632-83956	Sequence 83956, A
29	15.8	79.0	492	14	US-10-027-632-83957	Sequence 83957, A
30	15.8	79.0	492	14	US-10-027-632-298547	Sequence 298547,
31	15.8	79.0	492	14	US-10-027-632-298548	Sequence 298548,
32	15.8	79.0	840	12	US-10-191-803-211	Sequence 211, App
33	15.8	79.0	870	12	US-10-369-493-37126	Sequence 37126, A
c 34	15.8	79.0	1116	12	US-10-447-328-31	Sequence 31, Appl
c 35	15.8	79.0	1374	15	US-10-001-073-42	Sequence 42, Appl
c 36	15.8	79.0	1386	15	US-10-001-073-40	Sequence 40, Appl
c 37	15.8	79.0	1758	13	US-10-101-510-450	Sequence 450, App
c 38	15.8	79.0	2000	10	US-09-938-842A-4905	Sequence 4905, Ap
c 39	15.8	79.0	2000	12	US-09-938-842A-4905	Sequence 4905, Ap
c 40	15.8	79.0	2826	15	US-10-225-567A-43	Sequence 43, Appl
c 41	15.8	79.0	6456	13	US-10-311-455-980	Sequence 980, App
42	15.8	79.0	684707	12	US-10-398-221-9	Sequence 9, Appli
c 43	15.8	79.0	3011208	12	US-10-398-221-2058	Sequence 2058, Ap
c 44	15.8	79.0	3673778	13	US-10-312-841-1	Sequence 1, Appli
c 45	15.8	79.0	3673778	13	US-10-312-841-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-001-073-16
; Sequence 16, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
; APPLICANT: Liggett, Stephen
; APPLICANT: Small, Kirsten
; TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
; FILE REFERENCE: 13073-PCT
; CURRENT APPLICATION NUMBER: US/10/001,073
; CURRENT FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-001-073-16

Query Match 100.0%; Score 20; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db          1 AAAAACGCCAATGACCACAG 20
```

RESULT 2
US-09-825-923-1/c
; Sequence 1, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
; APPLICANT: Snapir, Amir
; APPLICANT: Heinonen, Paula
; APPLICANT: Alhopuro, Pia
; APPLICANT: Karvonen, Matti
; APPLICANT: Koulu, Markku

```
;  APPLICANT:  Pesonen, Ullamari
;  APPLICANT:  Scheinin, Mika
;  APPLICANT:  Salonen, Jukka T
;  APPLICANT:  Tuomainen, Tomi-Pekka
;  APPLICANT:  Lakka, Timo A
;  APPLICANT:  Nyyss"nen, Kristiina
;  APPLICANT:  Salonen, Riitta
;  APPLICANT:  Kauhanen, Jussi
;  APPLICANT:  Valkonen, Veli-Pekka
;  TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
;  TITLE OF INVENTION:  protein, and uses thereof
;  FILE REFERENCE: Alpha-2B-AR variant
;  CURRENT APPLICATION NUMBER: US/09/825,923
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/422,985
;  PRIOR FILING DATE: 2000-05-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1341)
;   OTHER INFORMATION: Coding sequence for variant human
;   OTHER INFORMATION: alpha-2B-adrenoceptor protein
US-09-825-923-1

  Query Match             100.0%;  Score 20;  DB 9;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 8.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1131 AAAAACGCCAATGACCACAG 1112


RESULT 3
US-10-077-870-1/c
; Sequence 1, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
;  APPLICANT: Salonen, Jukka T
;  TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
;  FILE REFERENCE: 0933-0183P
;  CURRENT APPLICATION NUMBER: US/10/077,870
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: FI 20010323
;  PRIOR FILING DATE: 2001-02-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1341)
;   OTHER INFORMATION: Coding sequence for variant human alpha-2B-adrenoceptor protein
US-10-077-870-1

  Query Match             100.0%;  Score 20;  DB 15;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 8.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1131 AAAAACGCCAATGACCACAG 1112


RESULT 4
US-10-001-073-2/c
; Sequence 2, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
;  APPLICANT: Liggett, Stephen
;  APPLICANT:  Small, Kirsten
;  TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
;  FILE REFERENCE: 13073-PCT
;  CURRENT APPLICATION NUMBER: US/10/001,073
;  CURRENT FILING DATE:  2001-11-01
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1344
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-001-073-2

  Query Match             100.0%;  Score 20;  DB 15;  Length 1344;
  Best Local Similarity   100.0%;  Pred. No. 8.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1131 AAAAACGCCAATGACCACAG 1112


RESULT 5
US-09-825-923-3/c
; Sequence 3, Application US/09825923
; Patent No. US20010016338A1
; GENERAL INFORMATION:
;  APPLICANT: Snapir, Amir
;  APPLICANT:  Heinonen, Paula
;  APPLICANT:  Alhopuro, Pia
;  APPLICANT:  Karvonen, Matti
;  APPLICANT:  Koulu, Markku
;  APPLICANT:  Pesonen, Ullamari
;  APPLICANT:  Scheinin, Mika
;  APPLICANT:  Salonen, Jukka T
;  APPLICANT:  Tuomainen, Tomi-Pekka
;  APPLICANT:  Lakka, Timo A
;  APPLICANT:  Nyyss"nen, Kristiina
;  APPLICANT:  Salonen, Riitta
;  APPLICANT:  Kauhanen, Jussi
;  APPLICANT:  Valkonen, Veli-Pekka
;  TITLE OF INVENTION: A DNA molecule encoding a variant alpha-2B-adrenoceptor
;  TITLE OF INVENTION:  protein, and uses thereof
;  FILE REFERENCE: Alpha-2B-AR variant
;  CURRENT APPLICATION NUMBER: US/09/825,923
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/422,985
;  PRIOR FILING DATE: 2000-05-25
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1350)
;   OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor
;   OTHER INFORMATION: protein
US-09-825-923-3

  Query Match             100.0%;  Score 20;  DB 9;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 8.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121
```

```
RESULT 6
US-10-077-870-3/c
; Sequence 3, Application US/10077870
; Publication No. US20030003470A1
; GENERAL INFORMATION:
;  APPLICANT: Salonen, Jukka T
;  TITLE OF INVENTION: Method for detecting a risk of hypertension and uses thereof
;  FILE REFERENCE: 0933-0183P
;  CURRENT APPLICATION NUMBER: US/10/077,870
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: FI 20010323
;  PRIOR FILING DATE: 2001-02-20
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1350)
;   OTHER INFORMATION: Coding sequence for human alpha-2B-adrenoceptor protein
US-10-077-870-3

  Query Match             100.0%;  Score 20;  DB 15;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 8.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121


RESULT 7
US-10-001-073-1/c
; Sequence 1, Application US/10001073
; Publication No. US20030113725A1
; GENERAL INFORMATION:
;  APPLICANT: Liggett, Stephen
;  APPLICANT:  Small, Kirsten
;  TITLE OF INVENTION: Alpha-2-adrenergic receptor polymorphisms
;  FILE REFERENCE: 13073-PCT
;  CURRENT APPLICATION NUMBER: US/10/001,073
;  CURRENT FILING DATE:  2001-11-01
;  NUMBER OF SEQ ID NOS: 53
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1353
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-001-073-1

  Query Match             100.0%;  Score 20;  DB 15;  Length 1353;
  Best Local Similarity   100.0%;  Pred. No. 8.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121


RESULT 8
US-10-305-720-1181/c
; Sequence 1181, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
;  APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
;  TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressio
;  FILE REFERENCE: PA-0002-1 CON
;  CURRENT APPLICATION NUMBER: US/10/305,720
;  CURRENT FILING DATE:  2002-11-26
;  PRIOR APPLICATION NUMBER: 09/016,434
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 1490
;  SOFTWARE: PERL Program
; SEQ ID NO 1181
;   LENGTH: 2072
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: GenBank ID No. US20040010136A1 g178197
US-10-305-720-1181

  Query Match             100.0%;  Score 20;  DB 12;  Length 2072;
  Best Local Similarity   100.0%;  Pred. No. 8.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1552 AAAAACGCCAATGACCACAG 1533


RESULT 9
US-10-225-567A-41/c
; Sequence 41, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
;  APPLICANT: LifeSpan Biosciences
;  APPLICANT:  Brown, Joseph P.
;  APPLICANT:  Burmer, Glenna C.
;  APPLICANT:  Roush, Christine L.
;  TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 1920-4-4
;  CURRENT APPLICATION NUMBER: US/10/225,567A
;  CURRENT FILING DATE:  2001-12-19
;  PRIOR APPLICATION NUMBER: 60/257,144
;  PRIOR FILING DATE: 2000-12-19
;  NUMBER OF SEQ ID NOS: 2292
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
;   LENGTH: 3274
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-225-567A-41

  Query Match             100.0%;  Score 20;  DB 15;  Length 3274;
  Best Local Similarity   100.0%;  Pred. No. 9.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db       1140 AAAAACGCCAATGACCACAG 1121


RESULT 10
US-09-770-445-297/c
; Sequence 297, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
;  APPLICANT:  Gorlach, Jorn
;  APPLICANT:  An, Yong-Qiang
;  APPLICANT:  Hamilton, Carol M.
;  APPLICANT:  Price, Jennifer L.
;  APPLICANT:  Raines, Tracy M.
;  APPLICANT:  Yu, Yang
;  APPLICANT:  Rameaka, Joshua G.
;  APPLICANT:  Page, Amy
;  APPLICANT:  Matthew, Abraham V.
;  APPLICANT:  Ledford, Brooke L.
;  APPLICANT:  Woessner, Jeffrey P.
;  APPLICANT:  Haas, William David
;  APPLICANT:  Garcia, Carlos A.
;  APPLICANT:  Kricker, Maja
;  APPLICANT:  Slader, Ted
```

```
;  APPLICANT:  Davis, Keith R.
;  APPLICANT:  Allen, Keith
;  APPLICANT:  Hoffman, Neil
;  APPLICANT:  Hurban, Patrick
;  TITLE OF INVENTION:  Expressed Sequences of Arabidopsis
;  TITLE OF INVENTION:  thaliana
;  FILE REFERENCE:  2023US (PARA-012PRV)
;  CURRENT APPLICATION NUMBER: US/09/770,445
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER:  US 60/178,472
;  PRIOR FILING DATE:  2000-01-27
;  NUMBER OF SEQ ID NOS:  999
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 297
;   LENGTH: 965
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(965)
;   OTHER INFORMATION: n = A,T,C or G
US-09-770-445-297

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 965;
  Best Local Similarity   94.7%;  Pred. No. 1.4e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACAG 20
              |||| ||||||||||||||
Db        219 AAAAGGCCAATGACCACAG 201


RESULT 11
US-10-369-493-42551/c
; Sequence 42551, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 42551
;   LENGTH: 1767
;   TYPE: DNA
;   ORGANISM: Anabaena PCC7120
US-10-369-493-42551

  Query Match             84.0%;  Score 16.8;  DB 12;  Length 1767;
  Best Local Similarity   90.0%;  Pred. No. 2.8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||||| |||||| ||||||
Db        923 AAAAACACCAATGGCCACAG 904


RESULT 12
US-10-398-221-1266
; Sequence 1266, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
;  APPLICANT: KUNST, Frederik
;  APPLICANT: GLASER, Philippe
;  TITLE OF INVENTION: Listeria innocua, genome and applications
;  FILE REFERENCE: 344 702 - US
;  CURRENT APPLICATION NUMBER: US/10/398,221
;  CURRENT FILING DATE:  2003-03-27
;  PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
;  PRIOR FILING DATE: 2001-10-04
;  PRIOR APPLICATION NUMBER: FR 00/12 697
;  PRIOR FILING DATE: 2000-10-04
;  NUMBER OF SEQ ID NOS: 4025
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1266
;   LENGTH: 788
;   TYPE: DNA
;   ORGANISM: Listeria monocytogenes-4B
US-10-398-221-1266

  Query Match             82.0%;  Score 16.4;  DB 12;  Length 788;
  Best Local Similarity   94.4%;  Pred. No. 4e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCAC 18
              |||||||||||||||| ||
Db        244 AAAAACGCCAATGACGAC 261


RESULT 13
US-09-738-626-743/c
; Sequence 743, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
;  APPLICANT: NAKAGAWA, SATOSHI
;  APPLICANT:  MIZOGUCHI, HIROSHI
;  APPLICANT:  ANDO, SEIKO
;  APPLICANT:  HAYASHI, MIKIRO
;  APPLICANT:  OCHIAI, KEIKO
;  APPLICANT:  YOKOI, HARUHIKO
;  APPLICANT:  TATEISHI, NAOKO
;  APPLICANT:  SENOH, AKIHIRO
;  APPLICANT:  IKEDA, MASATO
;  APPLICANT:  OZAKI, AKIO
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-125
;  CURRENT APPLICATION NUMBER: US/09/738,626
;  CURRENT FILING DATE:  2000-12-18
;  PRIOR APPLICATION NUMBER: JP 99/377484
;  PRIOR FILING DATE: 1999-12-16
;  PRIOR APPLICATION NUMBER: JP 00/159162
;  PRIOR FILING DATE: 2000-04-07
;  PRIOR APPLICATION NUMBER: JP 00/280988
;  PRIOR FILING DATE: 2000-08-03
;  NUMBER OF SEQ ID NOS: 7059
;  SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 743
;   LENGTH: 996
;   TYPE: DNA
;   ORGANISM: Corynebacterium glutamicum
US-09-738-626-743

  Query Match             82.0%;  Score 16.4;  DB 10;  Length 996;
  Best Local Similarity   94.4%;  Pred. No. 4.1e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 AACGCCAATGACCACAG 20
              ||| ||||||||||||||
Db        319 AAATGCCAATGACCACAG 302


RESULT 14
US-10-292-798-1745/c
; Sequence 1745, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
```

```
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;  FILE REFERENCE: 084335/166
;  CURRENT APPLICATION NUMBER: US/10/292,798
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: 10/017,161
;  PRIOR FILING DATE: 2001-12-18
;  PRIOR APPLICATION NUMBER: JP 2001-246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2070
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1745
;   LENGTH: 1282
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   LOCATION: source
;   FEATURE:
;   LOCATION: (1)..(1282)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(1082)
US-10-292-798-1745

  Query Match             82.0%;  Score 16.4;  DB 12;  Length 1282;
  Best Local Similarity   94.4%;  Pred. No. 4.2e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACA 19
              | ||||||||||||||||
Db       1071 ACAACGCCAATGACCACA 1054


RESULT 15
US-10-017-161-2099/c
; Sequence 2099, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
;  APPLICANT: SUWA, MAKIKO
;  APPLICANT:  ASAI, KIYOSHI
;  APPLICANT:  AKIYAMA, YUTAKA
;  APPLICANT:  ABURATANI, HIROYUKI
;  TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
;  FILE REFERENCE: 084335/0152
;  CURRENT APPLICATION NUMBER: US/10/017,161
;  CURRENT FILING DATE:  2002-12-18
;  PRIOR APPLICATION NUMBER: JP 2001/246789
;  PRIOR FILING DATE: 2001-06-18
;  NUMBER OF SEQ ID NOS: 2430
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2099
;   LENGTH: 1282
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: source
;   LOCATION: (1)..(1282)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (201)..(1082)
US-10-017-161-2099

  Query Match             82.0%;  Score 16.4;  DB 13;  Length 1282;
  Best Local Similarity   94.4%;  Pred. No. 4.2e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACA 19
              | ||||||||||||||||
Db       1071 ACAACGCCAATGACCACA 1054


Search completed: February 12, 2004, 06:16:00
Job time : 122.659 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 03:54:41 ; Search time 1081.71 Seconds
(without alignments)
449.373 Million cell updates/sec

Title: US-09-692-077D-16
Perfect score: 20
Sequence: 1 aaaaacgccaatgaccacag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	19	95.0	712	29	HGRA2BAR	AL606572 Shark alp
	2	18.4	92.0	986	10	BE788667	BE788667 601475536
c	3	17.4	87.0	527	9	AI993504	AI993504 701496517
	4	17.4	87.0	971	29	CNS07D83	AL440137 T3 end of
	5	17.4	87.0	1752	10	BF129496	BF129496 601811063
c	6	17	85.0	464	13	BY030666	BY030666 BY030666
	7	16.8	84.0	173	9	AI395866	AI395866 MA000059b
	8	16.8	84.0	179	28	AZ070845	AZ070845 RPCI-23-4
c	9	16.8	84.0	241	28	AQ106085	AQ106085 HS_3054_A
c	10	16.8	84.0	321	29	BZ664659	BZ664659 SALK_0996
c	11	16.8	84.0	393	10	BE239439	BE239439 EST392780
	12	16.8	84.0	424	28	AQ432032	AQ432032 HS_2270_A
c	13	16.8	84.0	436	9	AI735704	AI735704 at20d01.x
c	14	16.8	84.0	445	9	AI243573	AI243573 qu63c06.x
c	15	16.8	84.0	500	10	BG143944	BG143944 ut66h01.x
c	16	16.8	84.0	513	12	BM690995	BM690995 UI-E-CI1-
c	17	16.8	84.0	576	28	AZ217526	AZ217526 Sheared D
	18	16.8	84.0	635	13	BQ119722	BQ119722 EST605298
c	19	16.8	84.0	658	13	BQ109472	BQ109472 imageqc_7
	20	16.8	84.0	668	28	AZ352444	AZ352444 1M0090I17
	21	16.8	84.0	722	14	CB677776	CB677776 OSJNEe15B
	22	16.8	84.0	729	9	AI055134	AI055134 coau0003C
	23	16.8	84.0	742	14	CB630921	CB630921 OSIIEb08C
c	24	16.8	84.0	765	28	AZ907242	AZ907242 RPCI-24-1
c	25	16.8	84.0	788	29	DREA2CAR	AL606584 Zebrafish
c	26	16.8	84.0	855	13	BQ901034	BQ901034 AGENCOURT
c	27	16.8	84.0	858	28	BH270411	BH270411 CH230-19I
c	28	16.8	84.0	878	12	BI768763	BI768763 603058253
c	29	16.8	84.0	882	29	HIPA2C2AR	AL606568 Sea Horse
	30	16.8	84.0	889	29	CNS04XDZ	AL311552 Tetraodon
c	31	16.8	84.0	896	29	BZ203770	BZ203770 CH230-436
c	32	16.8	84.0	916	29	CNS01MR6	AL151283 Anopheles
	33	16.8	84.0	926	10	BG440575	BG440575 GA__Ea000
	34	16.8	84.0	1115	28	BH167326	BH167326 ENTTF94TF
	35	16.8	84.0	1336	12	BG910723	BG910723 602809750
c	36	16.4	82.0	143	10	BG549263	BG549263 947074A07
	37	16.4	82.0	355	9	AI859458	AI859458 wm12a08.x
c	38	16.4	82.0	405	10	BG549863	BG549863 947076C08
c	39	16.4	82.0	455	29	CC468063	CC468063 CH240_139
c	40	16.4	82.0	547	12	BM499110	BM499110 947041G05
c	41	16.4	82.0	593	9	AI619097	AI619097 486073C10
c	42	16.4	82.0	655	29	AG069974	AG069974 Pan trogl
	43	16.4	82.0	670	9	AW697748	AW697748 Str1-C6 S
c	44	16.4	82.0	679	29	BZ275203	BZ275203 CH230-392
	45	16.4	82.0	745	10	BG319648	BG319648 Zm03_06g0

ALIGNMENTS

RESULT 1
HGRA2BAR/c
LOCUS HGRA2BAR 712 bp DNA linear GSS 14-SEP-2001
DEFINITION Shark alpha2 adrenergic receptor gene fragment probably subtype b, genomic survey sequence.
ACCESSION AL606572
VERSION AL606572.1 GI:15591923
KEYWORDS GSS; Alpha2 adrenergic receptor gene.
SOURCE Hexanchus griseus
ORGANISM Hexanchus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; Elasmobranchii; Squalea; Notidanoidea; Hexanchiformes; Hexanchoidei; Hexanchidae; Hexanchus.
REFERENCE 1 (bases 1 to 712)
AUTHORS Hunter,C. and Elgar,G.
TITLE Alpha2 adrenergic receptor gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 712)
AUTHORS Hunter,C.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-2001) MRC Human Genome Mapping Project Resource Centre Hinxton, Cambridge, CB10 1SB. UK Email: biohelp@hgmp.mrc.ac.uk
FEATURES Location/Qualifiers
source 1. .712
/organism="Hexanchus griseus"
/mol_type="genomic DNA"

```
                     /db_xref="taxon:94987"
BASE COUNT      188 a    187 c    180 g    157 t
ORIGIN

  Query Match             95.0%;  Score 19;  DB 29;  Length 712;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACAG 20
              ||||||||||||||||||||
Db        683 AAAACGCCAATGACCACAG 665


RESULT 2
BE788667
LOCUS       BE788667                 986 bp    mRNA    linear   EST 20-OCT-2000
DEFINITION  601475536F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3878433 5',
            mRNA sequence.
ACCESSION   BE788667
VERSION     BE788667.1  GI:10209865
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 986)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP/Gazdar
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9642  row: g  column: 10
            High quality sequence stop: 254.
FEATURES             Location/Qualifiers
     source          1. .986
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3878433"
                     /tissue_type="large cell carcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_68"
                     /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.8 kb.   Library constructed by Life
                     Technologies. "
BASE COUNT      446 a    332 c    157 g     51 t
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 10;  Length 986;
  Best Local Similarity   95.0%;  Pred. No. 2.6e+02;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||||||||||||||||| |||
Db        280 AAAAACGCCAATGACCGCAG 299


RESULT 3
AI993504/c
LOCUS       AI993504                 527 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION  701496517 A. thaliana, Ohio State clone set Arabidopsis thaliana
            cDNA clone 701496517, mRNA sequence.
ACCESSION   AI993504
VERSION     AI993504.1  GI:5840409
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 527)
  AUTHORS   Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
            Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
            Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R.
            , Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carpio,T.,
            Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobriga,A., Murry,L.,
            Turner,C., Krikorian,S., Elder,L. and Hanson,D.
  TITLE     Arabidopsis thaliana Gene Expression MicroArray
  JOURNAL   Unpublished
COMMENT     Contact: David Smoller, Ph.D.
            Genome Systems, Inc., a wholly owned subsidiary of Incyte
            Pharmaceuticals, Inc.
            4633 World Parkway Circle, St. Louis, MO 63134, USA
            Tel: 877-577-2733
            Fax: 314-427-3324
            Email: service@genomesystems.com.
FEATURES             Location/Qualifiers
     source          1. .527
                     /organism="Arabidopsis thaliana"
                     /mol_type="mRNA"
                     /db_xref="taxon:3702"
                     /clone="701496517"
                     /clone_lib="A. thaliana, Ohio State clone set"
                     /note="cDNA library was made from selected clones from the
                     Arabidopsis thaliana Ohio  State clone set."
BASE COUNT      121 a    163 c    115 g    128 t
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 9;  Length 527;
  Best Local Similarity   94.7%;  Pred. No. 6.4e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACAG 20
              |||| ||||||||||||||||
Db        258 AAAAGGCCAATGACCACAG 240


RESULT 4
CNS07D83
LOCUS       CNS07D83                 971 bp    DNA     linear   GSS 08-JUL-2001
DEFINITION  T3 end of clone BD0AA010C06 of library BD0AA from strain CBS 94 of
            Candida tropicalis, genomic survey sequence.
ACCESSION   AL440137
VERSION     AL440137.1  GI:12223548
KEYWORDS    GSS.
SOURCE      Candida tropicalis
  ORGANISM  Candida tropicalis
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1  (bases 1 to 971)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 971)
  AUTHORS   Blandin,G., Ozier-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and
            Dujon,B.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 16. Candida
            tropicalis
  JOURNAL   FEBS Lett. 487 (1), 91-94 (2000)
  MEDLINE   20584726
```

```
   PUBMED   11152891
REFERENCE   3  (bases 1 to 971)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web :
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .971
                     /organism="Candida tropicalis"
                     /mol_type="genomic DNA"
                     /strain="CBS 94"
                     /db_xref="taxon:5482"
                     /clone="BD0AA010C06"
                     /clone_lib="BD0AA"
                     /note="end : T3"
BASE COUNT      343 a    166 c    149 g    312 t      1 others
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 29;  Length 971;
  Best Local Similarity   94.7%;  Pred. No. 8e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACA 19
              ||||||| ||||||||||
Db        818 AAAAACGACAATGACCACA 836


RESULT 5
BF129496
LOCUS       BF129496               1752 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  601811063R1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4054095 3',
            mRNA sequence.
ACCESSION   BF129496
VERSION     BF129496.1  GI:10968446
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1752)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM893  row: n  column: 16
            High quality sequence start: 132
            High quality sequence stop: 163.
FEATURES             Location/Qualifiers
     source          1. .1752
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4054095"
                     /tissue_type="primary B-cells from tonsils (cell line)"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_48"
                     /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by Ling
                     Hong in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      468 a    455 c    497 g    331 t      1 others
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 10;  Length 1752;
  Best Local Similarity   94.7%;  Pred. No. 9.9e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 AAAACGCCAATGACCACAG 20
              ||||||||||| ||||||
Db        251 AAAACGCCAATGGCCACAG 269


RESULT 6
BY030666/c
LOCUS       BY030666               464 bp    mRNA    linear   EST 06-DEC-2002
DEFINITION  BY030666 RIKEN full-length enriched, 1 cell embryo Mus musculus
            cDNA clone I0C0014E12 5', mRNA sequence.
ACCESSION   BY030666
VERSION     BY030666.1  GI:26136109
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 464)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H.
            , Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C.,
            Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A.,
            Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S.,
            Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani
            ,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest
            ,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A.
            , Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J.,
            Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M.,
            King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons
            ,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki
            ,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G.,
            Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D.,
            Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring
            ,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou
            ,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale
            ,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y.,
            Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa
            ,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A.,
            Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura
            ,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.
            , Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii
            ,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata
            ,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander
            ,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
```

Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hirozane ,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H., Miyazaki,A. , Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Sakai,K. , Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami ,M., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

```
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="I0C0014E12"
                     /cell_type="1 cell"
                     /dev_stage="1 cell embryo"
                     /clone_lib="RIKEN full-length enriched, 1 cell embryo"
BASE COUNT      144 a     83 c    126 g    111 t
ORIGIN

  Query Match             85.0%;  Score 17;  DB 13;  Length 464;
  Best Local Similarity   100.0%;  Pred. No. 9.6e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCA 17
              |||||||||||||||||
Db        428 AAAAACGCCAATGACCA 412


RESULT 7
AI395866
LOCUS       AI395866                173 bp    mRNA    linear   EST 11-FEB-1999
DEFINITION  MA000059b.FGR Adult worm cDNA expression library Schistosoma
            mansoni cDNA 5', mRNA sequence.
ACCESSION   AI395866
VERSION     AI395866.1  GI:4225413
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 173)
  AUTHORS   Oliveira,G.C. and Baba,J.
  TITLE     Cataloguing Schistosoma mansoni genes with expressed sequence tags
  JOURNAL   Unpublished
COMMENT     Contact: Oliveira, Guilherme
            Lab. Parasitologia Cel. e Mol.
            Centro de Pesquisas Rene Rachou - FIOCRUZ
            Av. Augusto de Lima 1715 Barro Preto, Belo Horizonte, MG, CEP 30190
            , Brazil
            Tel: 55 31 2953566
            Fax: 55 31 2952115
            Email: oliveira@netra.cpqrr.fiocruz.br
            Seq primer: M13 Reverse Universal Sequencing primer.
FEATURES             Location/Qualifiers
     source          1. .173
                     /organism="Schistosoma mansoni"
                     /mol_type="mRNA"
                     /strain="LE"
                     /db_xref="taxon:6183"
                     /sex="Mixed"
                     /dev_stage="adult"
                     /clone_lib="Adult worm cDNA expression library"
                     /note="Vector: Uni-Zap XR vector, Stratagene (pBluescript SK); Site_1: EcoRI; Site_2: XhoI; library was constructed and excised according to the manufacturer's instructions."
BASE COUNT       64 a     31 c     29 g     49 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 173;
  Best Local Similarity   90.0%;  Pred. No. 8.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||||||||||||||| | ||
Db        140 AAAAACGCCAATGACTAAAG 159


RESULT 8
AZ070845
LOCUS       AZ070845                179 bp    DNA     linear   GSS 30-MAR-2000
DEFINITION  RPCI-23-422G2.TV RPCI-23 Mus musculus genomic clone RPCI-23-422G2,
            genomic survey sequence.
ACCESSION   AZ070845
VERSION     AZ070845.1  GI:7362097
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 179)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret ,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished
COMMENT     Other_GSSs: RPCI-23-422G2.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 422  row: G  column: 2
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .179
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-422G2"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and  partially digested
```

with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). "
BASE COUNT 76 a 58 c 45 g 0 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 28; Length 179;
Best Local Similarity 90.0%; Pred. No. 8.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
|||||| |||| ||||||||
Db 78 AAAAACACCAAAGACCACAG 97

RESULT 9
AQ106085/c
LOCUS AQ106085 241 bp DNA linear GSS 28-AUG-1998
DEFINITION HS_3054_A1_E12_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3054 Col=23 Row=I, genomic survey sequence.
ACCESSION AQ106085
VERSION AQ106085.1 GI:3481441
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 241)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3054 row: I column: 23
Class: BAC ends
High quality sequence stop: 241.
FEATURES Location/Qualifiers
source 1. .241
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3054 Col=23 Row=I"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"
BASE COUNT 52 a 54 c 52 g 83 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 28; Length 241;
Best Local Similarity 90.0%; Pred. No. 9.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
| |||||||||||| |||||
Db 74 AGAAACGCCAATGAACACAG 55

RESULT 10
BZ664659/c
LOCUS BZ664659 321 bp DNA linear GSS 31-JAN-2003
DEFINITION SALK_099680.45.60.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_099680.45.60.x, genomic survey sequence.
ACCESSION BZ664659
VERSION BZ664659.1 GI:28181534
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 321)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P. , Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At2g23500.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .321
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_099680.45.60.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 91 a 40 c 86 g 104 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 29; Length 321;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
||| ||||||||| ||||||
Db 277 AAATACGCCAATGCCCACAG 258

RESULT 11
BE239439/c
LOCUS BE239439 393 bp mRNA linear EST 12-JUL-2000
DEFINITION EST392780 MHRP- Medicago truncatula cDNA clone pMHRP-47M19, mRNA sequence.
ACCESSION BE239439
VERSION BE239439.1 GI:9055622
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 393)
AUTHORS Harrison,M.J., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E., Cho,J. and Fraser,C.M.
TITLE ESTs from phosphate-starved roots of Medicago truncatula

```
  JOURNAL   Unpublished
COMMENT     Contact: Maria J. Harrison
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73401, USA
            Tel: 580-223-5810
            Fax: 580-221-7380
            Email: mjharrison@noble.org
            The Samuel Noble Roberts Foundation: N260974e
            TIGR sequence name:MTHAU82TK
            More information is available at. .
            http://chrysie.tamu.edu/medicago
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .393
                     /organism="Medicago truncatula"
                     /mol_type="mRNA"
                     /cultivar="A17"
                     /db_xref="taxon:3880"
                     /clone="pMHRP-47M19"
                     /tissue_type="roots"
                     /dev_stage="phosphate-starved"
                     /lab_host="XLOLR"
                     /clone_lib="MHRP-"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; At the trifoliate stage, M. truncatula plants were
                     transplanted to phosphate-free sand and grown for a
                     further 30 days. During this period, they were fertilized
                     twice weekly with 1/2 Hoaglands solutions containing 20uM
                     potassium phosphate. cDNA was prepared from polyA+
                     enriched RNA. The cDNA was directionally ligated into the
                     Unizap XR vector from Stratagene and packaged using
                     Gigapack III Gold packaging extracts. Plasmids containing
                     cDNA inserts were excised from the recombinant lambda-Zap
                     phage using Ex-assist helper phage and propagated in
                     XLOLR cells."
BASE COUNT       98 a     69 c     82 g    144 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 10;  Length 393;
  Best Local Similarity   90.0%;  Pred. No. 1.1e+03;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||||||||||| ||||||
Db        241 AAAAACGCCAATTGCCACAG 222


RESULT 12
AQ432032
LOCUS       AQ432032                424 bp    DNA     linear   GSS 31-MAR-1999
DEFINITION  HS_2270_A2_E11_T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2270 Col=22 Row=I, genomic survey
            sequence.
ACCESSION   AQ432032
VERSION     AQ432032.1  GI:4542367
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 424)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2270  row: I  column: 22
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 424.
FEATURES             Location/Qualifiers
     source          1. .424
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2270 Col=22 Row=I"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      120 a     87 c     81 g    129 t      7 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 28;  Length 424;
  Best Local Similarity   90.0%;  Pred. No. 1.2e+03;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 AAAAACGCCAATGACCACAG 20
              |||||  | ||||||||||||
Db        196 AAAAATGGCAATGACCACAG 215


RESULT 13
AI735704/c
LOCUS       AI735704                436 bp    mRNA    linear   EST 14-JUN-1999
DEFINITION  at20d01.x1 Barstead aorta HPLRB6 Homo sapiens cDNA clone
            IMAGE:2355649 3' similar to gb:M58458 40S RIBOSOMAL PROTEIN S4, X
            ISOFORM (HUMAN);, mRNA sequence.
ACCESSION   AI735704
VERSION     AI735704.1  GI:5057228
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 436)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .436
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2355649"
                     /sex="male"
                     /dev_stage="adult, age 64"
                     /lab_host="DH10B (phage resistant)"
                     /clone_lib="Barstead aorta HPLRB6"
                     /note="Organ: aorta; Vector: pT7T3D-Pac (Pharmacia) with a
```

modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCGGATCGAAC 3' and 5' GTTGGATCGG 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead."

BASE COUNT 128 a 103 c 122 g 81 t 2 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 436;
Best Local Similarity 90.0%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
||||| ||||||| ||||||
Db 96 AAAAATGCCAATGGCCACAG 77

RESULT 14
AI243573/c
LOCUS AI243573 445 bp mRNA linear EST 04-NOV-1998
DEFINITION qu63c06.x1 NCI_CGAP_Brn35 Homo sapiens cDNA clone IMAGE:1976746 3' similar to gb:M58458 40S RIBOSOMAL PROTEIN S4, X ISOFORM (HUMAN);, mRNA sequence.
ACCESSION AI243573
VERSION AI243573.1 GI:3838970
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 445)
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
unknown library type
Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .445
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1976746"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn35"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.33 kb. Tumor types include: meningioma, oligodendroglioma, astrocytoma (grade II), medulloblastoma, astrocytoma (grade IV). Life Technologies catalog #: 11544-012"
BASE COUNT 105 a 85 c 92 g 162 t 1 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 445;
Best Local Similarity 90.0%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
||||| ||||||||| ||||
Db 208 AAAAATGCCAATGACAACAG 189

RESULT 15
BG143944/c
LOCUS BG143944 500 bp mRNA linear EST 01-FEB-2001
DEFINITION ut66h01.x1 Soares_mouse_NMGB_bcell Mus musculus cDNA clone IMAGE:3332976 3', mRNA sequence.
ACCESSION BG143944
VERSION BG143944.1 GI:12647344
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 500)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Other_ESTs: ut66h01.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1077140
High quality sequence stop: 480.
FEATURES Location/Qualifiers
source 1. .500
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3332976"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares_mouse_NMGB_bcell"
/note="Organ: germinal B-cell; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCTGGTTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized; constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 169 a 88 c 72 g 171 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 500;
Best Local Similarity 90.0%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AAAAACGCCAATGACCACAG 20
|||| ||||||||||||||
Db 320 AAAACAGCCAATGACCACAG 301

Search completed: February 12, 2004, 06:06:09
Job time : 1092.71 secs